

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:13:39 ; Search time 80.65 Seconds
(without alignments)
9.547 Million cell updates/sec

Title: US-09-763-397A-7

Perfect score: 109

Sequence: 1 DIEKKICKMEKSCSVFNVVNS 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	109	100.0	397	1 CSP_PLAFO	P19597 plasmodium
2	109	100.0	412	1 CSP_PLAFA	P02893 plasmodium
3	109	100.0	424	1 CSP_PLAFT	P13814 plasmodium
4	109	100.0	442	1 CSP_PLAFW	P08307 plasmodium
5	107	98.2	388	1 CSP_PLARE	P26694 plasmodium
6	83	76.1	339	1 CSP_PLABE	P06915 plasmodium
7	83	76.1	347	1 CSP_PLABA	P23093 plasmodium
8	83	76.1	367	1 CSP_PLAYO	P06914 plasmodium
9	77	70.6	393	1 CSP_PLABR	P14593 plasmodium
10	77	70.6	429	1 CSP_PLAMA	P13815 plasmodium
11	71	65.1	378	1 CSP_PLACB	P08672 plasmodium
12	71	65.1	378	1 CSP_PLACL	P08675 plasmodium
13	71	65.1	398	1 CSP_PLACC	P08673 plasmodium
14	71	65.1	401	1 CSP_PLACG	P08674 plasmodium
15	71	65.1	419	1 CSP_PLACM	P08676 plasmodium
16	69	63.3	343	1 CSP_PLAVS	P13826 plasmodium
17	69	63.3	378	1 CSP_PLAVB	P08677 plasmodium
18	69	63.3	386	1 CSP_PLASI	Q03110 plasmodium
19	65	59.6	351	1 CSP_PLAKU	P04922 plasmodium
20	65	59.6	363	1 CSP_PLAKH	P02894 plasmodium
21	54	49.5	629	1 Z195_HUMAN	P14628 homo sapien
22	50	45.9	535	1 Z257_HUMAN	Q9V2Q1 homo sapien
23	46	42.2	536	1 ZF94_RAT	Q9V2K3 rattus norv
24	46	42.2	3866	1 HRX_MOUSE	P55200 mus musculus
25	46	42.2	3969	1 HRX_HUMAN	Q03164 homo sapien
26	44	40.4	327	1 RIR4_ASM2	P26713 african swi
27	44	40.4	803	1 ZN43_HUMAN	P17038 homo sapien
28	44	40.4	2339	1 RQC1_PLAFA	P27625 plasmodium
29	43.5	39.9	280	1 Y1BA_ECOLI	P24172 escherichia
30	43	39.4	273	1 ZN80_PANTR	P51506 pan troglod
31	43	39.4	365	1 GN2_DROME	P20478 drosophila
32	42	38.5	640	1 Y1F6_YEAST	P40522 saccharomyc
33	42	38.5	1038	1 ITA4_HUMAN	P13612 homo sapien

RESULT 1

ID	CSP_PLAFO	STANDARD;	PRT;	397 AA.
AC	P19597: Q25798;			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).			
OS	Plasmodium falciparum (isolate NF54).			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
OX	NCBI_TaxID=5843;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89345189; PubMed=2668895;			
RA	Campbell J.R.;			
RT	"DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate vaccine antigen."			
RL	Nucleic Acids Res. 17:5854-5854(1989).			
RN	[2]			
RP	REVISIONS.			
RA	Campbell J.R.;			
RL	Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92155298; PubMed=1346766;			
RA	Davis J.R., Cortese J.F., Herrington D.A., Murphy J.R., Clyde D.F., Thomas A.W., Baqar S., Cochran M.A., Thanassi J., Levine M.M., Hackett C.S.;			
RT	"Plasmodium falciparum: in vitro characterization and human infectivity of a cloned line."			
RL	Exp. Parasitol. 74:159-168(1992).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89364998; PubMed=2671723;			
RA	Caspers P., Gentz R., Matile H., Pink J.R., Sinigaglia F.;			
RT	"The circumsporozoite protein gene from NF54, a Plasmodium falciparum isolate used in malaria vaccine trials."			
RL	Mol. Biochem. Parasitol. 35:185-190(1989).			
CC	-!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOOROZITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).			
CC	-!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.			
CC	-!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; X15363; CAA33421.1; -.			

Q23467 caenorhabdi
P52385 human herpe
P03977 saccharomyc
P53923 saccharomyc
Q00194 b cmp-gate
Q99335 campylobact
Q05481 homo sapien
P13496 drosophila
Q02963 p genome po
P18247 p genome po
O51378 borrelia bu
P51502 cercopithec

ALIGNMENTS

34 41.5 38.1 664 1 TBX7_CAEEL
35 41.5 38.1 721 1 PRTP_HSV7J
36 41.5 37.6 378 1 AI3M_YEAST
37 41 37.6 493 1 YNL9_YEAST
38 41 37.6 690 1 CNG1_BOVIN
39 41 37.6 773 1 SYFB_CAMJE
40 41 37.6 1191 1 ZN91_HUMAN
41 41 37.6 1265 1 DYN4_DROME
42 41 37.6 3061 1 POLG_PVYHU
43 41 37.6 3063 1 POLG_PVYN
44 40 36.7 211 1 KAD_BORBU
45 40 36.7 213 1 ZN80_CERAE

DR EMBL; M83886; AAA29521.1; -.
DR EMBL; M22982; AAA29527.1; -.
DR PIR; S05428; S05428.
DR PIR; A45527; A45527.
DR InterPro; IPR003067; Crmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 397 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 101 272 43 X 4 AA TANDEM REPEATS OF N-A-N-P.
FT CONFLICT 194 194 A -> ANPNANPN (IN REF. 4).
SQ SEQUENCE 397 AA; 42646 MW; 9E81146F59EBCEA3 CRC64;

Query Match 100.0%; Score 109; DB 1; Length 397;
Best Local Similarity 100.0%; Pred. No. 5.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DIEKICKMEKCSSVFNVVNS 21
Db 363 DIEKICKMEKCSSVFNVVNS 383

RESULT 2
CSP_PLAFA STANDARD; PRT; 412 AA.
AC P02893;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84250215; PubMed=6204383;
RA Dime J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA Hochmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy E.P., Diggs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on
RT the sporozoite of the human malaria parasite Plasmodium falciparum.";
RL Science 225:593-599(1984).
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; K02194; AAA29524.1; -.
CC PIR; A03388; OZQOAF.
CC InterPro; IPR003067; Crmsprzoite.
CC InterPro; IPR000884; TSP1.
CC Pfam; PF00090; tsp.1; 1.
CC PRINTS; PR01303; CRMSPRZOITE.
CC SMART; SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 412 CIRCUMSPOROZOITE PROTEIN.

FT DOMAIN 123 290 41 X 4 AA TANDEM REPEATS OF P-N-A-N.
SQ SEQUENCE 412 AA; 44420 MW; 1EEED3DE9065F8 CRC64;

Query Match 100.0%; Score 109; DB 1; Length 412;
Best Local Similarity 100.0%; Pred. No. 5.7e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DIEKICKMEKCSSVFNVVNS 21
Db 378 DIEKICKMEKCSSVFNVVNS 398

RESULT 3
CSP_PLAFT STANDARD; PRT; 424 AA.
AC P13814;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium falciparum (isolate t4 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5846;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87315205; PubMed=3306373;
RA del Portillo H.A., Nussenzweig R.S., Enea V.;
RT "Circumsporozoite gene of a Plasmodium falciparum strain from
RT Thailand.";
RL Mol. Biochem. Parasitol. 24:289-294(1987).
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M19752; AAA29555.1; -.
CC PIR; A54533; A54533.
CC InterPro; IPR003067; Crmsprzoite.
CC InterPro; IPR000884; TSP1.
CC Pfam; PF00090; tsp.1; 1.
CC PRINTS; PR01303; CRMSPRZOITE.
CC SMART; SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 424 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 123 300 45 X 4 AA TANDEM REPEATS OF N-A-N-P.
SQ SEQUENCE 424 AA; 45610 MW; 710AB14238786CD9 CRC64;

Query Match 100.0%; Score 109; DB 1; Length 424;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DIEKICKMEKCSSVFNVVNS 21
Db 390 DIEKICKMEKCSSVFNVVNS 410

RESULT 4
CSP_PLAFT

```

RT malaria parasite Plasmodium falciparum." ;
RL J. Biol. Chem. 266:6686-6689(1991).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M60972; AAA29561.1; -
CC PIR; A39756; A39756.
CC InterPro; IPR003067; Crcmsprzoite.
CC InterPro; IPR000884; TSP1.
CC Pfam; PF000590; tsp_1; 1.
CC PRINTS; PR01303; CRCMSPRZOITE.
CC SMART; SM00209; TSPI; 1.
CC Malaria; Sporozoite; Repeat; Signal.
CC FT SIGNAL 1 16 PROBABLE.
CC FT CHAIN 17 388 CIRCUMSPOROZOITE PROTEIN.
CC FT DOMAIN 120 267 62 X 4 AA TANDEM REPEATS OF N-A-N-P.
CC FT SEQUENCE 388 AA; 42245 MW; C031EEFB2E35604 CRC64;
CC -----
Query Match 98.2%; Score 107; DB 1; Length 388;
Best Local Similarity 95.2%; Pred. No. 11e-08;
Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIEKKICKMEKCSSVFNVVNS 21
   1:|||||
Db 354 DLEKKICKMEKCSSVFNVVNS 374
   1:|||||

RESULT 6
CSP_PLABE
ID ID CSP_PLABE STANDARD; PRT; 339 AA.
AC P06915;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium berghei.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5821;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87089740; PubMed=2432395;
RA Eichinger D.J., Arnot D.E., Tam J.P., Nussenzweig V., Enea V.;
RT "Circumsporozoite protein of Plasmodium berghei: gene cloning and
RT identification of the immunodominant epitopes.";
RL Mol. Cell. Biol. 6:3965-3972(1986).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: M14135; AAA29577.1; -
 DR PIR: A25083; OZZQMB.
 DR InterPro: IPR003067; Crmsprzoite.
 DR InterPro: IPR000884; TSPL.
 DR Pfam: PF00090; tsp.1; 1.
 DR PRINTS: PR01303; CRCMSPRZOITE.
 DR SMART: SM00209; TSPL; 1.
 DR PROSITE: PS50092; TSPL; 1.
 DR Malaria; Sporozoite; Repeat; Signal.
 KW Malaria; Sporozoite; Repeat; Signal.
 FT SIGNAL 1 23 PROBABLE.
 FT CHAIN 24 339 CIRCUMSPOROZOITE PROTEIN.
 FT DOMAIN 93 196 13 X 8 AA TANDEM REPEATS.
 FT DOMAIN 206 238 16 X 2 AA TANDEM REPEATS OF P-Q.
 SQ SEQUENCE 339 AA; 37138 MW; E8068A6D1D9551B CRC64;

Query Match 76.1%; Score 83; DB 1; Length 339;
 Best Local Similarity 61.9%; Pred. No. 3.2e-05;
 Matches 13; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

OY 1 DIEKKICKMEKSSVNVNVS 21
 ||:|||||:|||||:|||||:|

DB 305 DIDTEICKMDKSSIFNVSN 325

RESULT 7

CSP_PLABA ID CSP_PLABA STANDARD; PRT; 347 AA.
 AC P23093;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
 OS Plasmodium berghei (strain Anka).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID:5823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE:90221834; PubMed:2193186;
 RA Lockyer M.J., Davies C.S., Suhrbier A., Sinden R.E.;
 RT "Nucleotide sequence of the Plasmodium berghei circumsporoite
 RT protein gene from the ANKA clone 2.34L.";
 RL Nucleic Acids Res. 18:376-376(1990).
 CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
 CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
 CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
 CC VERTEBRATE HOST).
 CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
 CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
 CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
 CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X17606; CAA35608.1; -
 DR PIR: S07873; OZZQMB.
 DR InterPro: IPR003067; Crmsprzoite.
 DR InterPro: IPR000884; TSPL.
 DR Pfam: PF00090; tsp.1; 1.
 DR PRINTS: PR01303; CRCMSPRZOITE.
 DR SMART: SM00209; TSPL; 1.
 DR PROSITE: PS50092; TSPL; 1.
 DR Malaria; Sporozoite; Repeat; Signal.

FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 347 CIRCUMSPOROZOITE PROTEIN.
 FT DOMAIN 93 204 13 X 8 AA REPEATS.
 FT DOMAIN 214 247 17 X 2 AA REPEATS OF P-Q.
 SQ SEQUENCE 347 AA; 37776 MW; 0EC240EE35681AF8 CRC64;

Query Match 76.1%; Score 83; DB 1; Length 347;
 Best Local Similarity 61.9%; Pred. No. 3.3e-05;
 Matches 13; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

OY 1 DIEKKICKMEKSSVNVNVS 21
 ||:|||||:|||||:|||||:|

DB 313 DIDTEICKMDKSSIFNVSN 333

RESULT 8

CSP_PLAYO ID CSP_PLAYO STANDARD; PRT; 367 AA.
 AC P06914;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
 OS Plasmodium berghei yoelii
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID:5862;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE:87137555; PubMed:3102479;
 RA Lal A.A., de la Cruz V.F., Welsh J.A., Charoenvit Y., Malay W.L.,
 RA McCutchan T.F.;
 RT "Structure of the gene encoding the circumsporoite protein of
 RT Plasmodium yoelii. A rodent model for examining antimalarial
 RT sporozoite vaccines.";
 RL J. Biol. Chem. 262:2937-2940(1987).
 RN [2]
 RP SEQUENCE OF 1-140 AND 260-367 FROM N.A.
 RX MEDLINE:88232798; PubMed:3287156;
 RA de la Cruz V.F., Lal A.A., McCutchan T.F.;
 RT "Variation among circumsporoite protein genes from rodent
 RT malarial.";
 RL Mol. Biochem. Parasitol. 28:31-38(1988).
 CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
 CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
 CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
 CC VERTEBRATE HOST).
 CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
 CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
 CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
 CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: J02695; AAA29558.1; -
 DR EMBL: M18821; AAA29559.1; -
 DR EMBL: M22698; AAA29560.1; -
 DR PIR: A26271; OZZQMB.
 DR InterPro: IPR003067; Crmsprzoite.
 DR InterPro: IPR000884; TSPL.
 DR Pfam: PF00090; tsp.1; 1.
 DR PRINTS: PR01303; CRCMSPRZOITE.
 DR SMART: SM00209; TSPL; 1.
 DR Malaria; Sporozoite; Repeat; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 367 CIRCUMSPOROZOITE PROTEIN
 FT DOMAIN 139 328 15 X 6 AA TANDEM REPEATS OF Q-G-P-G-A-P.


```

FT DOMAIN 229 260 8 X 4 AA TANDEM REPEATS OF Q-Q-P-P.
SQ SEQUENCE 367 AA; 38888 MW; 1EA56AFF7FFCB5E3 CRC64;

Query Match 76.1%; Score 83; DB 1; Length 367;
Best Local Similarity 61.9%; Pred. No. 3.5e-05;
Matches 13; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 DIEKKICKMEKCSSVFNVS 21
DB 333 DIDTEICKMDKCSSIFNVSN 353

RESULT 9
CSP_PLABR STANDARD; PRT; 393 AA.
ID CSP_PLABR AC
AC P14593;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (CS) (FRAGMENT).
OS Plasmodium brasilianum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5824;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8818654; PubMed=3128542;
RA Lal A.A., la Cruz V.F., Collins W.E., Procell P.M.,
RA McCutchan T.F.;
RT "Circumsporozoite protein gene from Plasmodium brasilianum. Animal
RT reservoirs for human malaria parasites?";
RL J. Biol. Chem. 263:5495-5498(1988).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J03203; AAA29553.1; -
CC PIR: A28615; A28615.
CC HSP: P04002; IWFA.
CC InterPro: IPR000884; TSP1.
CC Pfam: PF00090; tsp_1; 1.
CC PRINTS: PR01303; CIRCSPRZOITE.
CC SMART: SM00209; TSP1; 1.
CC PROSITE: PS50092; TSP1; 1.
CC SIGNAL 1 15
CC CHAIN 16 429 CIRCUMSPOROZOITE PROTEIN.
CC DOMAIN 110 314 4 AA TANDEM REPEATS.
CC SEQUENCE 393 AA; 35372 MW; B32944419BC600AA CRC64;

Query Match 70.6%; Score 77; DB 1; Length 393;
Best Local Similarity 57.1%; Pred. No. 0.00028;
Matches 12; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 DIEKKICKMEKCSSVFNVS 21
DB 359 DLETEICSLDKCSSIFNVSN 379

RESULT 10
CSP_PLAMA STANDARD; PRT; 378 AA.
ID CSP_PLAMA AC
AC P08672;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium cynomolgi (strain Berok).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5828;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87102878; PubMed=3802196;
RA Galinski M.R., Arnot D.E., Cochran A.H., Barnwell J.W.,

```

```

ID CSP_PLAMA STANDARD; PRT; 429 AA.
AC P13815;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium malariae.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5858;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89040027; PubMed=3054537;
RA Lal A.A., la Cruz V.F., Campbell G.H., Procell P.M., Collins W.E.,
RA McCutchan T.F.;
RT "Structure of the circumsporozoite gene of Plasmodium malariae.";
RL Mol. Biochem. Parasitol. 30:291-294(1988).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J03992; AAA29557.1; -
CC PIR: A54504; A54504.
CC HSP: P04002; IWFA.
CC InterPro: IPR003067; Crcmsprzoite.
CC InterPro: IPR000884; TSP1.
CC Pfam: PF00090; tsp_1; 1.
CC PRINTS: PR01303; CIRCSPRZOITE.
CC SMART: SM00209; TSP1; 1.
CC PROSITE: PS50092; TSP1; 1.
CC SIGNAL 1 15
CC CHAIN 16 429 CIRCUMSPOROZOITE PROTEIN.
CC DOMAIN 110 314 4 AA TANDEM REPEATS.
CC SEQUENCE 429 AA; 41596 MW; 3629D641D1C0BB7E CRC64;

Query Match 70.6%; Score 77; DB 1; Length 429;
Best Local Similarity 57.1%; Pred. No. 0.0003;
Matches 12; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 DIEKKICKMEKCSSVFNVS 21
DB 395 DLETEICSLDKCSSIFNVSN 415

RESULT 11
CSP_PLACB STANDARD; PRT; 378 AA.
ID CSP_PLACB AC
AC P08672;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium cynomolgi (strain Berok).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5828;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87102878; PubMed=3802196;
RA Galinski M.R., Arnot D.E., Cochran A.H., Barnwell J.W.,

```



```

Query Match      65.1%; Score 71; DB 1; Length 401;
Best Local Similarity 47.6%; Pred. No. 0.0022;
Matches 10; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY   1 DIEKICKMEKCSSVFNVVNS 21
DB   367 DLETEVTMDKCAGIFNVS 387
      | : | : | : ||| : 
-----
RESULT 15
CSP_PLACM          STANDARD;           PRT;    419 AA.
AC   P08676;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS   Plasmodium cynomolgi (strain Mulligan/NIH).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5832;
RN   [ ]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87102878; PubMed=3802196;
RA   Galinski M.R., Arnot D.E., Cochrane A.H., Barnwell J.W.,
RT   Nussenzeig R.S., Enea V.;
RF   "the circumsporozoite gene of the Plasmodium cynomolgi complex.";
RL   Cell 48:311-319(1987).
CC   CC -I- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC   CC -I- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC   CC -I- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
-----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
CC   EMBL; M15102; AAA29539.1; -.
DR   PIR; B26255; OZZQAM.
DR   InterPro; IPR003067; Circmsprozoite.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00090; tsp_1; 1.
DR   PRINTS; PR01303; CRCMSPROZOITE.
DR   SMART; SM00209; TSP1; 1.
KW   Malaria; Sporozoite; Repeat; Signal.
FT   SIGNAL              1..19 PROBABLE.
FT   CHAIN               20..419 CIRCUMSPOROZOITE PROTEIN.
FT   DOMAIN              99..314 54 X 4 AA TANDEM REPEATS OF N-A-[DG]-G.
SQ   SEQUENCE            419 AA; 38924 MW; 8F46CDD8A1B4EFF4 CRC64;

Query Match      65.1%; Score 71; DB 1; Length 419;
Best Local Similarity 47.6%; Pred. No. 0.0023;
Matches 10; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

Qy   1 DIEKICKMEKCSSVFNVVNS 21
Db   385 DLETEVTMDKCAGIFNVS 405
      | : | : | : ||| : 
-----
Search completed: January 29, 2002, 11:13:39
Job time: 815 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:12:10 ; Search time 285.36 seconds
(without alignments)
10.764 Million cell updates/sec

Title: US-09-763-397A-7

Perfect score: 109

Sequence: 1 DIEKKICKMEKCSSVFNVVNS 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_17:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	109	100.0	80	5	Q9U0Q4	Q9u0q4 plasmodium
2	109	100.0	80	5	Q9U0Q0	Q9u0q0 plasmodium
3	109	100.0	80	5	Q9U0P9	Q9u0p9 plasmodium
4	109	100.0	80	5	Q9U0P8	Q9u0p8 plasmodium
5	109	100.0	80	5	Q9U0P7	Q9u0p7 plasmodium
6	109	100.0	80	5	Q9U0P6	Q9u0p6 plasmodium
7	109	100.0	80	5	Q9U0P5	Q9u0p5 plasmodium
8	109	100.0	80	5	Q9U0P4	Q9u0p4 plasmodium
9	109	100.0	80	5	Q9U0P3	Q9u0p3 plasmodium
10	109	100.0	80	5	Q9U0P2	Q9u0p2 plasmodium
11	109	100.0	80	5	Q9U0P1	Q9u0p1 plasmodium
12	109	100.0	80	5	Q9TW97	Q9tw97 plasmodium
13	109	100.0	80	5	Q9TW83	Q9tw83 plasmodium
14	109	100.0	80	5	Q9TW76	Q9tw76 plasmodium
15	109	100.0	80	5	Q9TW01	Q9tw01 plasmodium
16	109	100.0	80	5	Q9TW00	Q9tw00 plasmodium
17	109	100.0	80	5	Q9TVQ0	Q9tvq0 plasmodium
18	109	100.0	80	5	Q9TVP9	Q9tvp9 plasmodium
19	109	100.0	80	5	Q9TVN9	Q9tvn9 plasmodium

20	109	100.0	115	5	Q25835	Q25835 plasmodium
21	109	100.0	115	5	Q25836	Q25836 plasmodium
22	109	100.0	115	5	Q25837	Q25837 plasmodium
23	109	100.0	115	5	Q25839	Q25839 plasmodium
24	109	100.0	115	5	Q9U934	Q9u934 plasmodium
25	109	100.0	115	5	Q9TVW7	Q9tvw7 plasmodium
26	109	100.0	117	5	Q25794	Q25794 plasmodium
27	109	100.0	117	5	Q25795	Q25795 plasmodium
28	109	100.0	117	5	Q25796	Q25796 plasmodium
29	109	100.0	117	5	Q25797	Q25797 plasmodium
30	109	100.0	393	5	Q99255	Q99255 plasmodium
31	109	100.0	408	5	Q25729	Q25729 plasmodium
32	109	100.0	416	5	Q25829	Q25829 plasmodium
33	109	100.0	420	5	Q25831	Q25831 plasmodium
34	109	100.0	420	5	Q25838	Q25838 plasmodium
35	109	100.0	424	5	Q27425	Q27425 plasmodium
36	109	100.0	424	5	Q99256	Q99256 plasmodium
37	109	100.0	432	5	Q27246	Q27246 plasmodium
38	109	100.0	432	5	Q25827	Q25827 plasmodium
39	109	100.0	436	5	Q27325	Q27325 plasmodium
40	109	100.0	436	5	Q25828	Q25828 plasmodium
41	109	100.0	442	5	Q25830	Q25830 plasmodium
42	109	100.0	452	5	Q25834	Q25834 plasmodium
43	106	97.2	80	5	Q9U0Q3	Q9u0q3 plasmodium
44	106	97.2	80	5	Q9U0Q1	Q9u0q1 plasmodium
45	105	96.3	79	5	Q9U0Q2	Q9u0q2 plasmodium

ALIGNMENTS

RESULT 1

Q9U0Q4

ID Q9U0Q4 PRELIMINARY; PRT; 80 AA.

AC Q9U0Q4;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).

GN CS.

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5833;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Bl;

RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;

RT "Sequence variation in the non-repeat region of the Plasmodium falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and Burmese field isolates and from laboratory strains.";

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ269941; CAB64167.1; .

DR InterPro; IPR000884; TSPL.

DR InterPro; IPR003067; Crmsprzoite.

DR Pfam; PF00090; tsp-1; 1.

DR PRINTS; PR01303; CRCMSPRZOITE.

DR SMART; SM00209; TSPL; 1.

FT NON_TER 1 80

FT NON_TER 80 80

SQ SEQUENCE 80 AA; 9102 MW; A3283B70CEE50FDE CRC64;

Query Match 100.0%; Score 109; DB 5; Length 80;

Best Local Similarity 100.0%; Pred. No. 8.9e-10;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIEKKICKMEKCSSVFNVVNS 21

Db 60 DIEKKICKMEKCSSVFNVVNS 80

RESULT 2

Q9U0Q0

```

ID Q9U0Q0 PRELIMINARY; PRT; 80 AA.
AC Q9U0Q0;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D28;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269948; CAB64173.1; -
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9046 MW; BA7689D18F031C3E CRC64;

Query Match 100.0%; Score 109; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.9e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIEKKICKMEKCSSVFNVNS 21
Db 60 DIEKKICKMEKCSSVFNVNS 80

RESULT 3
Q9U0P9 PRELIMINARY; PRT; 80 AA.
AC Q9U0P9;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=L1;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269951; CAB64176.1; -
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9032 MW; ADED6F0E266AD98E CRC64;

Query Match 100.0%; Score 109; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.9e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 DIEKKICKMEKCSSVFNVNS 21
Db 60 DIEKKICKMEKCSSVFNVNS 80

RESULT 4
Q9U0P8 PRELIMINARY; PRT; 80 AA.
AC Q9U0P8;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M4;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269957; CAB64237.1; -
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9073 MW; ABF404B8FB142B1E CRC64;

Query Match 100.0%; Score 109; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.9e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIEKKICKMEKCSSVFNVNS 21
Db 60 DIEKKICKMEKCSSVFNVNS 80

RESULT 5
Q9U0P7 PRELIMINARY; PRT; 80 AA.
AC Q9U0P7;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D4230;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269961; CAB64180.1; -
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80

```

.SQ SEQUENCE 80 AA; 9047 MW; BA769C90DB031C3E CRC64;
 Query Match 100.0%; Score 109; DB 5; Length 80;
 Best Local Similarity 100.0%; Pred. No. 8.9e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DIEKKICKMEKCSSVFNVVNS 21
 Db 60 DIEKKICKMEKCSSVFNVVNS 80
 RESULT 6
 Q9U0P6 PRELIMINARY; PRT; 80 AA.
 AC Q9U0P6; 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
 GN CS.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D4364;
 RA "de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
 RT "Sequence variation in the non-repeat region of the Plasmodium
 RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
 RT Burmese field isolates and from laboratory strains.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ269963; CAB64182.1; -
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; tsp_1; 1.
 DR PRINTS; PR01303; CRCMSPRZOITE.
 DR SMART; SM00209; TSP1; 1.
 FT NON_TER 1
 FT NON_TER 80
 SQ SEQUENCE 80 AA; 9002 MW; ICEAE08E6C9E976 CRC64;
 Query Match 100.0%; Score 109; DB 5; Length 80;
 Best Local Similarity 100.0%; Pred. No. 8.9e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DIEKKICKMEKCSSVFNVVNS 21
 Db 60 DIEKKICKMEKCSSVFNVVNS 80
 RESULT 7
 Q9U0P5 PRELIMINARY; PRT; 80 AA.
 AC Q9U0P5; 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
 GN CS.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D4363;
 RA "de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
 RT "Sequence variation in the non-repeat region of the Plasmodium
 RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
 RT Burmese field isolates and from laboratory strains.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ269963; CAB64184.1; -

DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; tsp_1; 1.
 DR PRINTS; PR01303; CRCMSPRZOITE.
 DR SMART; SM00209; TSP1; 1.
 FT NON_TER 1
 FT NON_TER 80
 SQ SEQUENCE 80 AA; 9000 MW; 03798BD18F0BF3B3 CRC64;
 Query Match 100.0%; Score 109; DB 5; Length 80;
 Best Local Similarity 100.0%; Pred. No. 8.9e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DIEKKICKMEKCSSVFNVVNS 21
 Db 60 DIEKKICKMEKCSSVFNVVNS 80
 RESULT 8
 Q9U0P4 PRELIMINARY; PRT; 80 AA.
 AC Q9U0P4; 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
 GN CS.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D4396;
 RA "de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
 RT "Sequence variation in the non-repeat region of the Plasmodium
 RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
 RT Burmese field isolates and from laboratory strains.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ269967; CAB64186.1; -
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; tsp_1; 1.
 DR PRINTS; PR01303; CRCMSPRZOITE.
 DR SMART; SM00209; TSP1; 1.
 FT NON_TER 1
 FT NON_TER 80
 SQ SEQUENCE 80 AA; 8988 MW; 0E7689D18F031B53 CRC64;
 Query Match 100.0%; Score 109; DB 5; Length 80;
 Best Local Similarity 100.0%; Pred. No. 8.9e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DIEKKICKMEKCSSVFNVVNS 21
 Db 60 DIEKKICKMEKCSSVFNVVNS 80
 RESULT 9
 Q9U0P3 PRELIMINARY; PRT; 80 AA.
 AC Q9U0P3; 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
 GN CS.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.

RESULT 13

Q9TW83 ID Q9TW83 PRELIMINARY; PRT; 80 AA.
AC Q9TW83;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PA17, D41, AND H1.
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ269952; CAB64177.1; -
DR EMBL: AJ269943; CAB64169.1; -
DR EMBL: AJ269944; CAB64174.1; -
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR003067; Crmsprzoite.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 8989 MW; 0E769C8AC1030149 CRC64;

Query Match 100.0%; Score 109; DB 5; Length 80;

Best Local Similarity 100.0%; Pred. No. 8.9e-10; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIEKKICKMEKCSSVFNVVNS 21
|||||
DB 60 DIEKKICKMEKCSSVFNVVNS 80

RESULT 14

Q9TW76 ID Q9TW76 PRELIMINARY; PRT; 80 AA.
AC Q9TW76;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PALO ALTO, AND D50;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ269954; CAB64179.1; -
DR EMBL: AJ269944; CAB64170.1; -
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR003067; Crmsprzoite.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9043 MW; 17282E319AE508B3 CRC64;

Query Match 100.0%; Score 109; DB 5; Length 80;

Best Local Similarity 100.0%; Pred. No. 8.9e-10; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIEKKICKMEKCSSVFNVVNS 21
|||||
DB 60 DIEKKICKMEKCSSVFNVVNS 80

Search completed: January 29, 2002, 11:12:10
Job time: 766 sec

Query Match 100.0%; Score 109; DB 5; Length 80;

Best Local Similarity 100.0%; Pred. No. 8.9e-10; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIEKKICKMEKCSSVFNVVNS 21
|||||
DB 60 DIEKKICKMEKCSSVFNVVNS 80

RESULT 15

Q9TW01 ID Q9TW01 PRELIMINARY; PRT; 80 AA.
AC Q9TW01;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M11, AND D7;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ269960; CAB64240.1; -
DR EMBL: AJ269947; CAB64172.1; -
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR003067; Crmsprzoite.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR PROSITE: PS50092; TSP1; 1.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9123 MW; 4614EEC68F0B1434 CRC64;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:21:43 ; Search time 310.82 Seconds
(without alignments)
2.145 Million cell updates/sec

Title: US-09-763-397A-8
Perfect score: 59
Sequence: 1 WSPCSVTCG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	9	12	AA13626
2	59	100.0	9	20	AAW97439
3	59	100.0	9	20	AAW81490
4	59	100.0	9	20	AAW81491
5	59	100.0	9	20	AAW81478
6	59	100.0	9	21	AAW70284
7	59	100.0	13	7	AAW60413
8	59	100.0	16	15	AAW51428
9	59	100.0	18	15	AAW51429
10	59	100.0	18	15	AAW51430
11	59	100.0	18	15	AAW51431

12	59	100.0	18	15	AAW51432
13	59	100.0	18	19	AAW59270
14	59	100.0	18	22	AAW49237
15	59	100.0	20	20	AAW03674
16	59	100.0	21	15	AAW51427
17	59	100.0	23	19	AAW59271
18	59	100.0	23	20	AAW22153
19	59	100.0	23	20	AAW22139
20	59	100.0	33	20	AAW22154
21	59	100.0	46	19	AAW59274
22	59	100.0	54	16	AAW71644
23	59	100.0	54	22	AAW48115
24	59	100.0	113	12	AAW13504
25	59	100.0	126	19	AAW59273
26	59	100.0	180	11	AAW07290
27	59	100.0	184	11	AAW07289
28	59	100.0	229	20	AAW03681
29	59	100.0	250	7	AAW60412
30	59	100.0	309	12	AAW13175
31	59	100.0	319	11	AAW07945
32	59	100.0	319	12	AAW13176
33	59	100.0	327	12	AAW13177
34	59	100.0	335	12	AAW13178
35	59	100.0	335	12	AAW13179
36	59	100.0	350	21	AAW70278
37	59	100.0	402	8	AAW70709
38	59	100.0	411	9	AAW83144
39	59	100.0	412	7	AAW60416
40	59	100.0	412	9	AAW80835
41	59	100.0	424	14	AAW37796
42	59	100.0	424	14	AAW37797
43	59	100.0	429	10	AAW90064
44	59	100.0	559	11	AAW05427
45	59	100.0	2150	21	AAW53898

ALIGNMENTS

RESULT 1

AA13626
ID AA13626 standard; peptide: 9 AA.

XX AA13626;

DT 31-OCT-1991 (first entry)

XX Thrombospondin peptide pl.

DE Antiviral agent; wound healing; platelet aggregation; thrombotic;

KW Thrombolytic.

XX Synthetic.

OS EP443404-A.

PN 28-AUG-1991.

PD 11-FEB-1991; 91EP-0101908.

XX 31-JAN-1991; 91US-0646531.

XX 22-FEB-1990; 90US-0483527.

XX (GRAC) GRACE W R & CO-CONN.

XX (MEDI-) MED COLLEGE OF PENNSYLVANIA.

XX Deutch AH, Tuszynski GP;

DR WPI; 1991-254044/35.

XX New peptide fragments and analogues of thrombospondin - useful

PT for inhibiting tumour metastasis, as clotting agents and to

PT promote or inhibit cell adhesion and immune modulation.

XX Claim 3; Page 26; 30pp; English.

XX The peptide may have an NH₂ gp on Gly(9) (peptide p6) and Trp(1)

CC may be formylated (peptide p5). The peptide is a synthetic

CC fragment of human thrombospondin and comprises the sequence motif

CC of Robson et al (Nature (1988) 335:79-82). It has thrombospondin-

CC like activity and can be used:

CC (1) to inhibit tumour cell metastasis and atherosclerosis;

CC (2) to promote or inhibit platelet aggregation, angiogenic activity,

CC thrombotic or thrombolytic activity, immune modulation and cell

CC adhesion;

CC (3) to promote wound healing; and

CC (4) as an antiviral agent (interferes with cell adhesion).

CC The peptide is prep'd. by std. synthesis techniques.

CC See also AAR13627-R13641.

XX Sequence 9 AA;

XX Query Match 100.0%; Score 59; DB 12; Length 9;

XX Best Local Similarity 100.0%; Pred. No. 4.3e+05;

XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSTVTCG 9

DB 1 wpscvstcg 9

|||||||

RESULT 2

AAW97439

ID AAW97439 standard; peptide; 9 AA.

AC AAW97439;

XX 19-MAY-1999 (first entry)

DT Peptide of the invention.

XX Neuronal sprouting; cerebral cortex; spinal cord cell; aggregation;

KW neurite defasciculation; regeneration; nervous system cell;

KW neurodegeneration; Alzheimer's; Parkinson's disease; multiple sclerosis;

KW myopathy; synapse formation; neuroblastoma.

XX Synthetic.

OS WO9903890-A1.

PN 28-JAN-1999.

XX 16-JUL-1998; 98WO-FR01556.

PF 16-JUL-1997; 97FR-0009016.

PR (UYAU-) UNIV AUVERGNE.

XX Gobron S, Meinzel A, Monnerie H;

PI WPI; 1999-132166/11.

DR New polypeptides for promoting growth of neurons - useful for

PT treatment of neurodegeneration, neuroblastoma and as additives for

PT neuronal cell cultures

XX Claim 1; Page 14; 29pp; French.

XX AAW97439-41 represent sequences that are not contained within a

CC peptide that stimulates neuronal (particularly axonal) sprouting

CC in neurons of the cerebral cortex, and in spinal cord cells inhibits

CC aggregation and defasciculation of neurites while increasing formation

CC of synaptic contacts. The peptide is used to regenerate nervous

CC system cells, particularly for treating neurodegeneration

CC (e.g. Alzheimer's or Parkinson's diseases, multiple sclerosis and

CC myopathy), other conditions requiring regeneration (particularly

CC elongation and synapse formation) or neuroblastoma. The peptide

CC can also be used as an additive for neuronal cell cultures.

XX Sequence 9 AA;

XX Query Match 100.0%; Score 59; DB 20; Length 9;

XX Best Local Similarity 100.0%; Pred. No. 4.3e+05;

XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSTVTCG 9

DB 1 wpscvstcg 9

|||||||

RESULT 3

AAW81490

ID AAW81490 standard; peptide; 9 AA.

XX AAW81490;

XX 29-JAN-1999 (first entry)

DT Thrombospondin synthetic analog compound p5.

XX Thrombospondin; thrombin sensitive protein; TSP; cell-adhesion;

KW mitogenic activity; chemotactic; haemostatic; tumour; microbial;

KW parasite; metastasis; platelet aggregation; fibrinolytic; malaria;

KW immune modulation; wound healing; atherosclerosis; angiogenesis;

KW complement modulator; diagnostic reagent.

XX Synthetic.

OS US5840692-A.

PN 24-NOV-1998.

XX 07-JUN-1995; 95US-0488273.

XX 07-JUN-1995; 95US-0488273.

PR 22-FEB-1996; 90US-0483527.

PR 03-JUN-1994; 92US-0896527.

PR 20-AUG-1993; 93US-0110146.

PR 19-DEC-1994; 94US-0359263.

XX (GRAC) GRACE & CO-CONN W R.

PA (UYAL-) UNIV ALLEGHENY HEALTH SCI.

XX Deutch AH, Tuszynski GP;

PI WPI; 1999-034078/03.

DR Method for mimicking or inhibiting thrombospondin activity - using

PT thrombospondin peptides

XX Claim 3; Column 29; 19pp; English.

XX Sequences AAW81478 to AAW81491 represent synthetic peptide analogs of

CC thrombospondin (thrombin sensitive protein or TSP) that retain

CC thrombospondin-like activity. The invention provides a method for (a)

CC mimicking a biological activity of thrombospondin; (b) promoting

CC thrombotic activity; or (c) inhibiting a biological activity of

CC thrombospondin other than thrombotic activity. The method comprises

CC administering any of these peptides. Biological activities of

CC thrombospondin include cell adhesion-promoting activity, cell mitogenic

CC activity, cell chemotactic activities, haemostatic activities, and

CC activities that derive from these activities, e.g. tumour-cell,
CC microbial or parasite metastasis activity, platelet aggregating
CC activity, fibrinolytic activity and immune modulation. The peptides are
CC capable of inhibiting tumour metastasis. The peptides are useful in wound
CC healing, atherosclerosis, malaria, thrombotic and thrombolytic conditions
CC and angiogenesis, and as cell attachment promoters, complement modulators
CC and diagnostic reagents.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 59; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPSCVTCG 9
 |||||||
DB 1 wpscvtcg 9

RESULT 4
AAW81491
ID AAW81491 standard; peptide; 9 AA.
AC AAW81491;
DT 29-JAN-1999 (first entry)
DE Thrombospondin synthetic analog compound p6.
KW Thrombospondin; thrombin sensitive protein; TSP; cell-adhesion;
KW mitogenic activity; chemotactic; haemostatic; tumour; microbial;
KW parasite; metastasis; platelet aggregation; fibrinolytic; malaria;
KW immune modulation; wound healing; atherosclerosis; angiogenesis;
KW complement modulator; diagnostic reagent.
OS Synthetic.
FH Key Location/Qualifiers
FT Modified-site 9 /note= "C-terminal amide"
FT
PW US5840692-A.
PD 24-NOV-1998.
PP 07-JUN-1995; 95US-0488273.
PR 07-JUN-1995; 95US-0488273.
PR 22-FEB-1990; 90US-0483527.
PR 09-JUN-1992; 92US-0896527.
PR 20-AUG-1993; 93US-0110146.
PR 19-DEC-1994; 94US-0359263.
XX (GRAC) GRACE & CO-CONN W R.
PA (UYAL-) UNIV ALLEGHENY HEALTH SCI.
XX Deutch AH, Tuszynski GP;
XX WPI: 1999-034078/03.
XX Method for mimicking or inhibiting thrombospondin activity - using
PT thrombospondin peptides
PS Claim 3; Column 29; 19pp; English.
XX Sequences AAW81478 to AAW81491 represent synthetic peptide analogs of
CC thrombospondin (thrombin sensitive protein or TSP) that retain
CC thrombospondin-like activity. The invention provides a method for (a)
CC mimicking a biological activity of thrombospondin; (b) promoting
CC thrombotic activity; or (c) inhibiting a biological activity of
CC thrombospondin other than thrombotic activity. The method comprises
CC administering any of these peptides. Biological activities of
CC thrombospondin include cell adhesion-promoting activity, cell mitogenic
CC activity, cell chemotactic activities, haemostatic activities, and

CC thrombospondin include cell adhesion-promoting activity, cell mitogenic
CC activity, cell chemotactic activities, haemostatic activities, and
CC activities that derive from these activities, e.g. tumour-cell,
CC microbial or parasite metastasis activity, platelet aggregating
CC activity, fibrinolytic activity and immune modulation. The peptides are
CC capable of inhibiting tumour metastasis. The peptides are useful in wound
CC healing, atherosclerosis, malaria, thrombotic and thrombolytic conditions
CC and angiogenesis, and as cell attachment promoters, complement modulators
CC and diagnostic reagents.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 59; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPSCVTCG 9
 |||||||
DB 1 wpscvtcg 9

RESULT 5
AAW81478
ID AAW81478 standard; peptide; 9 AA.
AC AAW81478;
DT 29-JAN-1999 (first entry)
DE Thrombospondin synthetic analog compound p1.
KW Thrombospondin; thrombin sensitive protein; TSP; cell-adhesion;
KW mitogenic activity; chemotactic; haemostatic; tumour; microbial;
KW parasite; metastasis; platelet aggregation; fibrinolytic; malaria;
KW immune modulation; wound healing; atherosclerosis; angiogenesis;
KW complement modulator; diagnostic reagent.
OS Synthetic.
FH Key Location/Qualifiers
FT Modified-site 9 /note= "C-terminal amide"
FT
PW US5840692-A.
PD 24-NOV-1998.
PP 07-JUN-1995; 95US-0488273.
PR 07-JUN-1995; 95US-0488273.
PR 22-FEB-1990; 90US-0483527.
PR 09-JUN-1992; 92US-0896527.
PR 20-AUG-1993; 93US-0110146.
PR 19-DEC-1994; 94US-0359263.
XX (GRAC) GRACE & CO-CONN W R.
PA (UYAL-) UNIV ALLEGHENY HEALTH SCI.
XX Deutch AH, Tuszynski GP;
XX WPI: 1999-034078/03.
XX Method for mimicking or inhibiting thrombospondin activity - using
PT thrombospondin peptides
PS Claim 3; Column 29; 19pp; English.
XX Sequences AAW81478 to AAW81491 represent synthetic peptide analogs of
CC thrombospondin (thrombin sensitive protein or TSP) that retain
CC thrombospondin-like activity. The invention provides a method for (a)
CC mimicking a biological activity of thrombospondin; (b) promoting
CC thrombotic activity; or (c) inhibiting a biological activity of
CC thrombospondin other than thrombotic activity. The method comprises
CC administering any of these peptides. Biological activities of
CC thrombospondin include cell adhesion-promoting activity, cell mitogenic
CC activity, cell chemotactic activities, haemostatic activities, and

CC activities that derive from these activities, e.g. tumour-cell,
CC microbial or parasite metastasis activity, platelet aggregating
CC activity, fibrinolytic activity and immune modulation. The peptides are
CC capable of inhibiting tumour metastasis. The peptides are useful in wound
CC healing, atherosclerosis, malaria, thrombotic and thrombolytic conditions
CC and angiogenesis, and as cell attachment promoters, complement modulators
CC and diagnostic reagents.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 59; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
Db 1 wpscvtcg 9

RESULT 6

AAY70284
ID AAY70284 standard; peptide; 9 AA.

XX
AC AAY70284;

XX
DT 06-JUN-2000 (first entry)

XX
DE Plasmodium falciparum SSP-2 antigenic epitope, P547.

XX
KW Recombinant protein; CDC/NIAIDVAC-1; multivalent; malaria; vaccine;
KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
KW Pf27; antiparasitic; prevention; anti-CDC/NIAIDVAC-1 antibody.

XX
OS Plasmodium falciparum.

XX
PN WO200011179-A1.

XX
PD 02-MAR-2000.

XX
PF 19-AUG-1998; 99WO-US18869.

XX
PR 21-AUG-1998; 98US-0097703.

XX
PA (NIAID-) NAT INST IMMUNOLOGY.

XX
PA (SSH) US DEPT HEALTH & HUMAN SERVICES.

XX
PI Lal AA, Shi YP, Hasnain SE;

XX
DR WPI: 2000-237654/20.

XX
PT Novel recombinant protein as vaccine for treating malarial infection
PT comprises antigenic peptides obtained from different stages of
PT plasmodium falciparum life cycle.

XX
PS Claim 2; Page 16; 52pp; English.

XX
CC The present sequence is the antigenic epitope P547, derived from
CC sporozoite surface protein-2 (SSP-2) of the sporozoite stage of
CC Plasmodium falciparum. It is used in the construction of recombinant
CC protein CDC/NIAIDVAC-1, which is a multivalent, multistage malarial
CC vaccine. The recombinant protein comprises, melittin signal peptide,
CC (His)6 tag, T-cell epitope from tetanus toxoid and 21 antigenic epitopes
CC from circumsporozoite protein (CSP), sporozoite surface protein-2
CC (SSP-2), liver stage antigen-1 (LSA-1), merozoite surface protein-1
CC (MSP-1), MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding
CC antigen-175 (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete
CC specific antigen, Pf27. These epitopes were obtained at different stages
CC of the life cycle of P. falciparum. CDC/NIAIDVAC-1 vaccine has

CC antiparasitic activity and can be used for treatment and prevention of
CC malarial infections. Anti-CDC/NIAIDVAC-1 antibodies can be used for
CC detecting P. falciparum in biological samples.

XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 59; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
Db 1 wpscvtcg 9

RESULT 7

AAP60413
ID AAP60413 standard; Protein; 13 AA.

XX
AC AAP60413;

XX
DT 13-JUN-1991 (first entry)

XX
DE Synthetic peptide fragment antigenic for the malaria parasite.

XX
KW Sporozite; vaccination.

XX
OS Plasmodium falciparum.

XX
PN EPI66410-A.

XX
PD 02-JAN-1986.

XX
PF 24-JUN-1985; 85EP-0107794.

XX
PR 26-JUN-1984; 84US-0624564.

XX
PA (USDC) US SEC OF COMMERCE.

XX
PA (USGO) US GOVERNMENT.

XX
PA (USSA) US SEC OF THE ARMY.

XX
PI McCutchan TF, Dame JB, Williams JL, Schneider I;

XX
DR WPI: 1986-008635/02.

XX
PT New immunologically active pure synthetic peptide(s) - used for
PT protection against infection by malaria parasite.

XX
PS Claim 11; Page 42; 49pp; English.

XX
CC The peptides may be synthesised in pure form and used to generate an
CC immune response in vaccination against malaria. The featured repeat
CC units are claimed and must be present in copies of 2-1000, in the
CC P.falciparum genome, the first is coded for 37 times.

XX
CC See also AAM60362.

XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 59; DB 7; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
Db 3 wpscvtcg 11

RESULT 8

AAR51428
ID AAR51428 standard; peptide; 16 AA.

XX

```

AC AAR51428;
XX
XX 26-OCT-1994 (first entry)
XX
XX Circumsporozoite protein region II+ peptide #2.
XX
XX Region II+; circumsporozoite; CS; Plasmodium; inhibitor; binding;
KW receptor; hepatocytes; malaria-susceptible; mammal; infection;
KW antibody; liver; human hepatocyte cell line; HepG2.
XX
XX Plasmodium vivax.
XX
XX WO9406464-A.
XX
XX 31-MAR-1994.
XX
XX 17-SEP-1993; 93WO-US08800.
XX
XX 17-SEP-1992; 92US-0947033.
XX
XX (UYNV ) UNIV NEW YORK STATE.
XX
XX Cerami C, Frevert U, Nussenzweig V, Sinnis P;
XX
XX WPI; 1994-118161/14.
XX
XX Peptide(s) corresp. to Region II+ of circumsporozoite - are used
PT for inhibiting hepatocyte invasion by malarial sporozoites for
PT preventing malaria infection
XX
XX Claim 8; Page 32; 42pp; English.
XX
XX The sequences given in AAR51427-35 are peptide fragments derived from
CC region II+ of the circumsporozoite (CS) protein derived from various
CC Plasmodium species. These peptides fragments may be used as an
CC inhibitor for the binding of CS polypeptides to receptors of
CC hepatocytes from malaria-susceptible mammals. These peptides can be
CC administered to malaria-susceptible mammals to prevent infection,
CC and they can be used to produce antibodies which can be used to
CC prevent infection. The peptides were tested for inhibition of CS
CC binding to liver sections or to the human hepatocyte cell line HepG2.
XX
XX Sequence 16 AA;

Query Match 100.0%; Score 59; DB 15; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTGC 9
Db | | | | | | | |
2 wpcsvtcg 10

RESULT 9
AAR51429
ID AAR51429 standard; peptide; 18 AA.
XX
XX AAR51429;
XX
XX 26-OCT-1994 (first entry)
XX
XX Circumsporozoite protein region II+ peptide #3.
XX
XX Region II+; circumsporozoite; CS; Plasmodium; inhibitor; binding;
KW receptor; hepatocytes; malaria-susceptible; mammal; infection;
KW antibody; liver; human hepatocyte cell line; HepG2.
XX
XX Plasmodium malariae.
XX
XX WO9406464-A.
XX
XX 31-MAR-1994.

```

```

XX
XX 17-SEP-1993; 93WO-US08800.
XX
XX 17-SEP-1992; 92US-0947033.
XX
XX (UYNV ) UNIV NEW YORK STATE.
XX
XX Cerami C, Frevert U, Nussenzweig V, Sinnis P;
XX
XX WPI; 1994-118161/14.
XX
XX Peptide(s) corresp. to Region II+ of circumsporozoite - are used
PT for inhibiting hepatocyte invasion by malarial sporozoites for
PT preventing malaria infection
XX
XX Claim 8; Page 32; 42pp; English.
XX
XX The sequences given in AAR51427-35 are peptide fragments derived from
CC region II+ of the circumsporozoite (CS) protein derived from various
CC Plasmodium species. These peptides fragments may be used as an
CC inhibitor for the binding of CS polypeptides to receptors of
CC hepatocytes from malaria-susceptible mammals. These peptides can be
CC administered to malaria-susceptible mammals to prevent infection,
CC and they can be used to produce antibodies which can be used to
CC prevent infection. The peptides were tested for inhibition of CS
CC binding to liver sections or to the human hepatocyte cell line HepG2.
XX
XX Sequence 18 AA;

Query Match 100.0%; Score 59; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTGC 9
Db | | | | | | | |
2 wpcsvtcg 10

RESULT 10
AAR51430
ID AAR51430 standard; peptide; 18 AA.
XX
XX AAR51430;
XX
XX 26-OCT-1994 (first entry)
XX
XX Circumsporozoite protein region II+ peptide #4.
XX
XX Region II+; circumsporozoite; CS; Plasmodium; inhibitor; binding;
KW receptor; hepatocytes; malaria-susceptible; mammal; infection;
KW antibody; liver; human hepatocyte cell line; HepG2.
XX
XX Plasmodium knowlesi.
XX
XX WO9406464-A.
XX
XX 31-MAR-1994.
XX
XX 17-SEP-1993; 93WO-US08800.
XX
XX 17-SEP-1992; 92US-0947033.
XX
XX (UYNV ) UNIV NEW YORK STATE.
XX
XX Cerami C, Frevert U, Nussenzweig V, Sinnis P;
XX
XX WPI; 1994-118161/14.
XX
XX Peptide(s) corresp. to Region II+ of circumsporozoite - are used
PT for inhibiting hepatocyte invasion by malarial sporozoites for
PT preventing malaria infection
XX
XX

```

PS Claim 8; Page 32; 42pp; English.

XX The sequences given in AAR51427-35 are peptide fragments derived from
CC region II+ of the circumsporozoite (CS) protein derived from various
CC Plasmodium species. These peptides fragments may be used as an
CC inhibitor for the binding of CS polypeptides to receptors of
CC hepatocytes from malaria-susceptible mammals. These peptides can be
CC administered to malaria-susceptible mammals to prevent infection,
CC and they can be used to produce antibodies which can be used to
CC prevent infection. The peptides were tested for inhibition of CS
CC binding to liver sections or to the human hepatocyte cell line HepG2.

XX Sequence 18 AA;

Query Match 100.0%; Score 59; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTGC 9
Db 2 wpscvtcg 10

RESULT 11

AAR51431
ID AAR51431 standard; peptide; 18 AA.

XX AAR51431;

DT 26-OCT-1994 (first entry)

DE Circumsporozoite protein region II+ peptide #5.

KW Region II+; circumsporozoite; CS; Plasmodium; inhibitor; binding;
KW receptor; hepatocytes; malaria-susceptible; mammal; infection;
KW antibody; liver; human hepatocyte cell line; HepG2.

XX Plasmodium cynomolgi.

XX WO9406464-A.

XX 31-MAR-1994.

XX 17-SEP-1993; 93WO-US08800.

XX 17-SEP-1992; 92US-0947033.

XX (UYNV) UNIV NEW YORK STATE.

XX Cerami C, Frevert U, Nussenzweig V, Sinnis P;

XX WPI; 1994-118161/14.

XX Peptide(s) corresp. to Region II+ of circumsporozoite - are used
PT for inhibiting hepatocyte invasion by malarial sporozoites for
PT preventing malaria infection

PS Claim 8; Page 32; 42pp; English.

XX The sequences given in AAR51427-35 are peptide fragments derived from
CC region II+ of the circumsporozoite (CS) protein derived from various
CC Plasmodium species. These peptides fragments may be used as an
CC inhibitor for the binding of CS polypeptides to receptors of
CC hepatocytes from malaria-susceptible mammals. These peptides can be
CC administered to malaria-susceptible mammals to prevent infection,
CC and they can be used to produce antibodies which can be used to
CC prevent infection. The peptides were tested for inhibition of CS
CC binding to liver sections or to the human hepatocyte cell line HepG2.

XX Sequence 18 AA;

XX

Query Match 100.0%; Score 59; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTGC 9
Db 2 wpscvtcg 10

RESULT 12

AAR51432
ID AAR51432 standard; peptide; 18 AA.

XX AAR51432;

XX 26-OCT-1994 (first entry)

DE Circumsporozoite protein region II+ peptide #6.

KW Region II+; circumsporozoite; CS; Plasmodium; inhibitor; binding;
KW receptor; hepatocytes; malaria-susceptible; mammal; infection;
KW antibody; liver; human hepatocyte cell line; HepG2.

XX Plasmodium brasilianum.

XX WO9406464-A.

XX 31-MAR-1994.

XX 17-SEP-1993; 93WO-US08800.

XX 17-SEP-1992; 92US-0947033.

XX (UYNV) UNIV NEW YORK STATE.

XX Cerami C, Frevert U, Nussenzweig V, Sinnis P;

XX WPI; 1994-118161/14.

XX Peptide(s) corresp. to Region II+ of circumsporozoite - are used
PT for inhibiting hepatocyte invasion by malarial sporozoites for
PT preventing malaria infection

PS Claim 8; Page 32; 42pp; English.

XX The sequences given in AAR51427-35 are peptide fragments derived from
CC region II+ of the circumsporozoite (CS) protein derived from various
CC Plasmodium species. These peptides fragments may be used as an
CC inhibitor for the binding of CS polypeptides to receptors of
CC hepatocytes from malaria-susceptible mammals. These peptides can be
CC administered to malaria-susceptible mammals to prevent infection,
CC and they can be used to produce antibodies which can be used to
CC prevent infection. The peptides were tested for inhibition of CS
CC binding to liver sections or to the human hepatocyte cell line HepG2.

XX Sequence 18 AA;

Query Match 100.0%; Score 59; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTGC 9
Db 2 wpscvtcg 10

RESULT 13

AAW59270
ID AAW59270 standard; peptide; 18 AA.

XX AAW59270;

XX

DT 27-AUG-1998 (first entry)
 DE P. falciparum circumsporozoite region II peptide fragment.
 KW Malaria; gene therapy; hepatocyte; liver; circumsporozoite; ligand;
 KW targeted delivery; therapy; disease; cancer; hepatitis; cystic fibrosis;
 KW hypercholesterolaemia; phenylketonuria; haemophilia.
 XX Plasmodium falciparum.
 XX US5766899-A.
 PN 16-JUN-1998.
 PD 27-FEB-1995; 95US-0395602.
 PF 27-FEB-1995; 95US-0395602.
 PR 27-FEB-1995; 95US-0395602.
 XX (TEXA) UNIV TEXAS SYSTEM.
 PA Ding Z, Kuo MT;
 XX WPI; 1998-361692/31.
 DR Complexes for targetted delivery of nucleic acids to hepatocytes -
 PT containing Plasmodium circumsporozoite polypeptide as targetting
 PT ligand
 XX Claim 2; Column 29-30; 34pp; English.
 PS AAM59270-W59274 are fragments of malarial circumsporozoite (CS) region
 CC II isolated from Plasmodium falciparum. These fragments can be used as
 CC ligands in a method for the targeted delivery of nucleic acid to cells
 CC in culture or cells in vivo, especially where the cells are hepatocytes.
 CC Therapy of diseases such as cancer, malaria, hepatitis, cystic fibrosis,
 CC hypercholesterolaemia, phenylketonuria and haemophilia is mentioned.
 CC CS polypeptides are liver cell specific with rapid hepatic invasion.
 CC They are more efficient than the prior art asialoorosomucoid (ASOR)
 CC ligands, of which there may be an accumulation in certain diseases due to
 CC receptor downregulation.
 XX Sequence 18 AA;
 SQ

Query Match 100.0%; Score 59; DB 19; Length 18;
 Best Local Similarity 100.0%; Pred. No. 0.017;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSPCSVTCG 9
 DB 2 wpscvtcg 10

RESULT 14
 AAB49237
 ID AAB49237 standard; Protein; 18 AA.
 XX
 AC AAB49237;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 DE Peptide RII derived from malaria circumsporozoite surface protein.
 XX
 KW Adenoviral vector; inverted terminal repeat; ITR; gene therapy;
 KW sickle cell anemia; thalassemia; atherosclerosis; restenosis;
 KW cancer.
 XX Synthetic.
 OS
 XX WO200073478-A2.
 PN
 XX 07-DEC-2000.
 PD
 XX

PF 01-JUN-2000; 2000WO-US15442.
 XX
 PR 01-JUN-1999; 99US-0137213.
 PR 22-OCT-1999; 99US-0161097.
 XX
 PA (UNIW) UNIV WASHINGTON.
 XX
 PI Lieber A, Shayakhmetov D, Farrar D, Papayannopoulou T;
 XX WPI; 2001-049942/06.
 DR
 XX Recombinant adenoviral vector containing transgene, used to produce
 PT gutless vectors for gene therapy, targetable to selected cells and
 PT lacking antigenicity -
 XX
 PS Example 2; Page 97; 156pp; English.
 XX
 CC The present invention relates to a first generation recombinant
 CC adenoviral vector, part of which integrates into a host cell genome. Two
 CC inverted terminal repeats (ITRs) allow integration of a transgene into
 CC the host genome. The invention can be used to make 'gutless' vectors
 CC for gene therapy, e.g. of sickle cell anemia or thalassemia (targeting
 CC hematopoietic cells) or atherosclerosis or restenosis (targeting
 CC endothelial cells), or more generally a wide range of genetic diseases,
 CC cancers and infectious diseases.
 XX
 SQ Sequence 18 AA;
 XX

Query Match 100.0%; Score 59; DB 22; Length 18;
 Best Local Similarity 100.0%; Pred. No. 0.017;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSPCSVTCG 9
 DB 2 wpscvtcg 10

RESULT 15
 AAY03674
 ID AAY03674 standard; peptide; 20 AA.
 XX
 AC AAY03674;
 XX
 DT 07-JUN-1999 (first entry)
 XX
 DE Amino acid sequence of the malaria (M) string CTL epitope TRAP AM.
 XX
 KW CD8+ T-cell; immune response; antigen; priming composition; CTL; epitope;
 KW cytotoxic T lymphocyte; boosting; poxvirus vector; PV; pathogen; tumour;
 KW malaria; parasite; P. falciparum; viral; bacterial; parasitic; cancer;
 KW melanoma; HIV; breast; colon; vaccination.
 XX
 OS Plasmodium falciparum.
 XX
 PN WO9856919-A2.
 XX
 PD 17-DEC-1998.
 XX
 PF 09-JUN-1998; 98WO-GB01681.
 XX
 PR 09-JUN-1997; 97GB-0011957.
 XX
 XX (ISIS-) ISIS INNOVATION LTD.
 PA
 XX Blanchard T, Gilbert SC, Hanke T, Hill AVS, McMichael AJ;
 PI Plebanski M, Schneider J, Smith GL;
 XX WPI; 1999-070325/06.
 DR N-PSDB; AAX29214.
 XX
 PT Generating CD8-positive T cell response to target antigen using
 PT recombinant poxvirus - for treating or preventing malaria and HIV

PT infection, also epitope strings from Plasmodium and HIV

XX Claim 38; Page 19; 85pp; English.

XX The invention relates to methods and reagents for generating a
CC protective CD8+ T-cell immune response against at least one target
CC antigen. The kits of the invention comprises (i) as priming composition,
CC a source of one or more CD8+ T-cell (cytotoxic T lymphocytes-(CTL))
CC epitopes of the target antigen, plus a carrier and (ii) as boosting
CC composition a source of CTL epitopes, with at least one CTL epitope the
CC same as used in (i), with this source being a non-replicating or
CC replication-impaired recombinant poxvirus vector (PVV) plus a carrier. If
CC the source of CTL epitopes in (i) is a viral vector, then the vector in
CC (ii) is from a different virus. The kits are used to generate an immune
CC response (prophylactic or therapeutic) against pathogens or tumours,
CC specifically against malaria parasites such as P. falciparum, or HIV, and
CC also many other bacterial, viral or parasitic pathogens. The kits are
CC also used for protective response against melanoma and cancer of breast
CC or colon, and generally wherever a strong CD8+ response is protective.
CC The boosting composition may be used alone to boost a naturally primed
CC response against malaria. The specified PVV provide an excellent booster
CC effect, better than that from wild-type poxvirus, resulting in complete
CC rather than partial protection against sporozoite challenge. Also PVV are
CC safer to use than wild-type virus. Sequences AAY03661-680 represent CTL
CC peptide epitopes of the malaria (M) string.

XX Sequence 20 AA:

Query Match 100.0%; Score 59; DB 20; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTTCG 9
DB 3 wpscsvtcgg 11
|||||

Search completed: January 29, 2002, 10:21:43
Job time: 419 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:24:03 : Search time 133.18 Seconds
(without alignments)
1.521 Million cell updates/sec

Title: US-09-763-397A-8
Perfect score: 59
Sequence: 1 WSPCSVTCG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	9	1	US-07-646-531D-1
2	59	100.0	9	1	US-07-646-531D-15
3	59	100.0	9	1	US-07-646-531D-16
4	59	100.0	9	2	US-08-488-273-1
5	59	100.0	9	2	US-08-488-273-13
6	59	100.0	9	2	US-08-488-273-14
7	59	100.0	9	2	US-08-858-971-1
8	59	100.0	9	6	5426100-1
9	59	100.0	18	1	US-08-395-602A-1
10	59	100.0	18	2	US-08-021-625D-1
11	59	100.0	23	1	US-08-395-602A-2
12	59	100.0	23	2	US-08-021-625D-2
13	59	100.0	23	4	US-08-986-659B-9
14	59	100.0	23	4	US-08-986-659B-33
15	59	100.0	33	4	US-08-986-659B-34
16	59	100.0	46	1	US-08-395-602A-5
17	59	100.0	46	2	US-08-021-625D-5
18	59	100.0	114	1	US-08-309-60A-2
19	59	100.0	126	1	US-08-395-602A-4
20	59	100.0	126	1	US-08-021-625D-4
21	59	100.0	412	1	US-08-313-288B-18
22	59	100.0	423	1	US-08-760-797A-1
23	59	100.0	424	2	US-08-760-797A-3
24	59	100.0	424	4	US-08-932-929B-1
25	59	100.0	424	4	US-08-932-929B-3
26	59	100.0	559	1	US-08-313-288B-14
27	56	94.9	9	1	US-07-646-531D-11

28	56	94.9	9	2	US-08-488-273-11	Sequence 11, Appl
29	56	94.9	378	6	5171843-9	Patent No. 5171843
30	55	93.2	53	2	US-08-799-173A-9	Sequence 9, Appl1
31	55	93.2	56	1	US-07-862-021B-16	Sequence 16, Appl
32	55	93.2	56	5	PCT-US93-03164-16	Sequence 16, Appl
33	55	93.2	392	2	US-08-799-173A-7	Sequence 7, Appl1
34	55	93.2	802	1	US-07-862-021B-12	Sequence 12, Appl
35	55	93.2	802	1	US-08-313-288B-12	Sequence 12, Appl
36	55	93.2	802	5	PCT-US93-03164-12	Sequence 12, Appl
37	55	93.2	807	1	US-07-862-021B-10	Sequence 10, Appl
38	55	93.2	807	1	US-08-313-288B-10	Sequence 10, Appl
39	55	93.2	807	5	PCT-US93-03164-10	Sequence 10, Appl
40	53	89.8	9	1	US-07-646-531D-8	Sequence 8, Appl1
41	53	89.8	9	2	5426100-8	Sequence 8, Appl1
42	53	89.8	9	6	5426100-8	Sequence 8, Appl1
43	51	86.4	9	1	US-07-646-531D-3	Sequence 3, Appl1
44	51	86.4	9	1	US-07-646-531D-9	Sequence 9, Appl1
45	51	86.4	9	1	US-07-646-531D-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-07-646-531D-1
; Sequence 1, Application US/07646531D
; Patent No. 5200397
; GENERAL INFORMATION:
; APPLICANT: Deutch, Alan Howard
; TITLE OF INVENTION: Peptide Fragments and Analogs of Thrombospondin
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. R. Grace & Co.-Conn.
; STREET: 7379 Route 32
; CITY: Columbia
; STATE: Maryland
; COUNTRY: USA
; ZIP: 21044
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/646,531D
; FILING DATE: 19910131
; ATTORNEY/AGENT INFORMATION:
; NAME: Appleby, Vanessa L.
; REGISTRATION NUMBER: 33223
; REFERENCE/DOCKET NUMBER: 01-7896
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 531-4515
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-646-531D-1

Query Match 100.0%; Score 59; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
|||||||
Db 1 WSPCSVTCG 9

RESULT 2

```
US-07-646-531D-15
; Sequence 15, Application US/07646531D
; Patent No. 5200397
; GENERAL INFORMATION:
; APPLICANT: Deutch, Alan Howard
; APPLICANT: Tuszyński, George Paul
; TITLE OF INVENTION: Peptide Fragments and Analogs of
; TITLE OF INVENTION: Thrombospondin
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. R. Grace & Co.-Conn.
; STREET: 7379 Route 32
; CITY: Columbia
; STATE: Maryland
; COUNTRY: USA
; ZIP: 21044
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/646,531D
; FILING DATE: 19910131
; ATTORNEY/AGENT INFORMATION:
; NAME: Appleby, Vanessa L.
; REGISTRATION NUMBER: 33223
; REFERENCE/DOCKET NUMBER: 01-7896
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 531-4515
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: the tryptophan at position 1 is n-
; OTHER INFORMATION: formyl-tryptophan and the glycine at position 9 is carboxyami
US-07-646-531D-15

Query Match 100.0%; Score 59; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTCG 9
Db 1 WSPCSVTCG 9

RESULT 3
US-07-646-531D-16
; Sequence 16, Application US/07646531D
; Patent No. 5200397
; GENERAL INFORMATION:
; APPLICANT: Deutch, Alan Howard
; APPLICANT: Tuszyński, George Paul
; TITLE OF INVENTION: Peptide Fragments and Analogs of
; TITLE OF INVENTION: Thrombospondin
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. R. Grace & Co.-Conn.
; STREET: 7379 Route 32
; CITY: Columbia
; STATE: Maryland
; COUNTRY: USA
; ZIP: 21044
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: Word Perfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/646,531D
; FILING DATE: 19910131
; ATTORNEY/AGENT INFORMATION:
; NAME: Appleby, Vanessa L.
; REGISTRATION NUMBER: 33223
; REFERENCE/DOCKET NUMBER: 01-7896
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 531-4515
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: the glycine at position 9 is
; OTHER INFORMATION: carboxyamide glycine
US-07-646-531D-16

Query Match 100.0%; Score 59; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTCG 9
Db 1 WSPCSVTCG 9

RESULT 4
US-08-488-273-1
; Sequence 1, Application US/08488273
; Patent No. 5840692
; GENERAL INFORMATION:
; APPLICANT: Deutch, Alan H.
; APPLICANT: Tuszyński, George P.
; TITLE OF INVENTION: PEPTIDE FRAGMENTS AND ANALOGS OF
; TITLE OF INVENTION: THROMBOSPONDIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
; STREET: 1601 Market Street, 36th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103-2398
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,273
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/359,263
; FILING DATE: 19-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/110,146
; FILING DATE: 20-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/896,527
; FILING DATE: 09-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/483,527
; FILING DATE: 22-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Leary Ph.D., Kathryn
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: 9598-204
```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 567-2020
TELEFAX: (215) 567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRAINEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-488-273-1

Query Match 100.0%; Score 59; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTGC 9
DB 1 WSPCSVTGC 9

RESULT 5

US-08-488-273-13
Sequence 13, Application US/08488273
Patent No. 5840592

GENERAL INFORMATION:
APPLICANT: Deutch, Alan H.
TITLE OF INVENTION: PEPTIDE FRAGMENTS AND ANALOGS OF
THROMBOSPONDIN
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: 1601 Market Street, 36th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103-2398

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/488.273
APPLICATION NUMBER: US/08/488.273
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/359,263
FILING DATE: 19-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,146
FILING DATE: 20-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,527
FILING DATE: 09-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/483,527
FILING DATE: 22-FEB-1990

ATTORNEY/AGENT INFORMATION:
NAME: Leary Ph.D., Kathryn
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9598-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 567-2020
TELEFAX: (215) 567-2991
TELEX: 831-494

INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid

STRAINEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1
OTHER INFORMATION: the tryptophan at position
OTHER INFORMATION: 1 is n-formyl-tryptophan.
FEATURE:
NAME/KEY: Peptide
LOCATION: 9
OTHER INFORMATION: the glycine at position 9
OTHER INFORMATION: is carboxamide.
US-08-488-273-13

Query Match 100.0%; Score 59; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTGC 9
DB 1 WSPCSVTGC 9

RESULT 6

US-08-488-273-14
Sequence 14, Application US/08488273
Patent No. 5840592

GENERAL INFORMATION:
APPLICANT: Deutch, Alan H.
TITLE OF INVENTION: PEPTIDE FRAGMENTS AND ANALOGS OF
THROMBOSPONDIN
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: 1601 Market Street, 36th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103-2398

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/488.273
APPLICATION NUMBER: US/08/488.273
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/359,263
FILING DATE: 19-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,146
FILING DATE: 20-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,527
FILING DATE: 09-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/483,527
FILING DATE: 22-FEB-1990

ATTORNEY/AGENT INFORMATION:
NAME: Leary Ph.D., Kathryn
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9598-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 567-2020
TELEFAX: (215) 567-2991
TELEX: 831-494

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 9
OTHER INFORMATION: the glycine at position 9
OTHER INFORMATION: is carboxyamide glycine."
US-08-488-273-14

Query Match 100.0%; Score 59; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WSPCSVTCG 9
Db 1 WSPCSVTCG 9

RESULT 7
US-08-858-971-1
; Sequence 1, Application US/08858971
; Patent No. 5886142
; GENERAL INFORMATION:
; APPLICANT: M. Thakur, V. Pallela
; TITLE OF INVENTION: Radiolabeled Thrombus Imaging Agents
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 210 Lake Drive East, Suite 201
; CITY: Cherry Hill
; STATE: NJ
; COUNTRY: USA
; ZIP: 08002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM 486
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,971
; FILING DATE: herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: JEFF-0196
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 779-2400
; TELEFAX: (609) 779-8488
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-858-971-1

Query Match 100.0%; Score 59; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WSPCSVTCG 9
Db 1 WSPCSVTCG 9

RESULT 8
5426100-1
; Patent No. 5426100
; APPLICANT: DEUTCH, ALAN H.;TUSZYNSKI, GEORGE
; TITLE OF INVENTION: PEPTIDE FRAGMENTS AND ANALOGS OF
; THROMBOSPONDIN
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,146
; FILING DATE: 20-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 896,527
; FILING DATE: 09-JUN-1992
; APPLICATION NUMBER: 483,527
; FILING DATE: 22-FEB-1990
; SEQ ID NO: 1:
; LENGTH: 9
5426100-1

Query Match 100.0%; Score 59; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WSPCSVTCG 9
Db 1 WSPCSVTCG 9

RESULT 9
US-08-395-602A-1
; Sequence 1, Application US/08395602A
; Patent No. 5766899
; GENERAL INFORMATION:
; APPLICANT: Kuo, M. Tien
; APPLICANT: Ding, Zhi-Ming
; TITLE OF INVENTION: Targeted Nucleic Acid Delivery into
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/395,602A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSC:410/HYL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-395-602A-1

Query Match 100.0%; Score 59; DB 1; Length 18;

Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTCG 9
Db 2 WSPCSVTCG 10

RESULT 10

US-08-021-625D-1
; Sequence 1, Application US/08021625D
; Patent No. 5976851
; GENERAL INFORMATION:
; APPLICANT: Kuo, M. Tien
; TITLE OF INVENTION: Targeted Nucleic Acid Delivery into
; TITLE OF INVENTION: Liver Cells
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/021,625D
; FILING DATE: 16-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSC:410/HYL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-021-625D-1

Query Match 100.0%; Score 59; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTCG 9
Db 2 WSPCSVTCG 10

RESULT 11

US-08-395-602A-2
; Sequence 2, Application US/08395602A
; Patent No. 5766899
; GENERAL INFORMATION:
; APPLICANT: Kuo, M. Tien
; TITLE OF INVENTION: Targeted Nucleic Acid Delivery into
; TITLE OF INVENTION: Liver Cells
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433

; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/395,602A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSC:410/HYL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-395-602A-2

Query Match 100.0%; Score 59; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTCG 9
Db 2 WSPCSVTCG 10

RESULT 12

US-08-021-625D-2
; Sequence 2, Application US/08021625D
; Patent No. 5976851
; GENERAL INFORMATION:
; APPLICANT: Kuo, M. Tien
; TITLE OF INVENTION: Targeted Nucleic Acid Delivery into
; TITLE OF INVENTION: Liver Cells
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/021,625D
; FILING DATE: 16-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSC:410/HYL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-021-625D-2

Query Match 100.0%; Score 59; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
DB 2 WSPCSVTCG 10

RESULT 13
US-08-986-659B-9
Sequence 9, Application US/08986659B
Patent No. 6171591
GENERAL INFORMATION:
APPLICANT: Hall, Stephen G.
TITLE OF INVENTION: RECOMBINANT NODAVIRUS RELATED
COMPOSITIONS AND METHODS
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Olson & Hierl, Ltd.
STREET: 20 No. 6171591th Wacker Drive, 36th Floor
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/986,659B
FILING DATE: 08-DEC-1997
CLASSIFICATION: 424
PRIOR APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Tallivaldis Cepuritis
REGISTRATION NUMBER: 20,818
REFERENCE/DOCKET NUMBER: 549.0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-580-1180
TELEFAX: 312-580-1189
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-986-659B-9

Query Match 100.0%; Score 59; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
DB 2 WSPCSVTCG 10

RESULT 14

US-08-986-659B-33
Sequence 33, Application US/08986659B
Patent No. 6171591
GENERAL INFORMATION:
APPLICANT: Hall, Stephen G.
TITLE OF INVENTION: RECOMBINANT NODAVIRUS RELATED
COMPOSITIONS AND METHODS
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Olson & Hierl, Ltd.
STREET: 20 No. 6171591th Wacker Drive, 36th Floor
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/986,659B
FILING DATE: 08-DEC-1997
CLASSIFICATION: 424
PRIOR APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Tallivaldis Cepuritis
REGISTRATION NUMBER: 20,818
REFERENCE/DOCKET NUMBER: 549.0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-580-1180
TELEFAX: 312-580-1189
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-986-659B-33

Query Match 100.0%; Score 59; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
DB 2 WSPCSVTCG 10

RESULT 15
US-08-986-659B-34
Sequence 34, Application US/08986659B
Patent No. 6171591
GENERAL INFORMATION:
APPLICANT: Hall, Stephen G.
TITLE OF INVENTION: RECOMBINANT NODAVIRUS RELATED
COMPOSITIONS AND METHODS
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Olson & Hierl, Ltd.
STREET: 20 No. 6171591th Wacker Drive, 36th Floor
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/986.659B
FILING DATE: 08-DEC-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Talivaldis Cepuritis
REGISTRATION NUMBER: 20,818
REFERENCE/DOCKET NUMBER: 549.0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-580-1180
TELEFAX: 312-580-1189
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-986-659B-34

Query Match 100.0%; Score 59; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTCG 9
| | | | | | | | | |
Db 7 WSPCSVTCG 15

Search completed: January 29, 2002, 10:24:03
Job time: 509 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:26:36 ; Search time 144.96 seconds
(without alignments)
4.729 Million cell updates/sec

Title: US-09-763-397A-8

Perfect score: 59

Sequence: 1 WSPCSVTCG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	378	1 OZ2QAB	circumsporozoite p
2	59	100.0	378	1 OZ2QAL	circumsporozoite p
3	59	100.0	388	2 A39756	circumsporozoite p
4	59	100.0	398	1 OZ2QAS	circumsporozoite p
5	59	100.0	401	1 OZ2QAC	circumsporozoite p
6	59	100.0	405	2 S05428	circumsporozoite p
7	59	100.0	412	1 OZ2QAF	circumsporozoite p
8	59	100.0	419	1 OZ2QAM	circumsporozoite p
9	59	100.0	424	2 A54533	circumsporozoite p
10	59	100.0	429	2 A54504	circumsporozoite p
11	59	100.0	442	2 A54529	circumsporozoite p
12	59	100.0	485	2 A60610	circumsporozoite p
13	59	100.0	559	2 S04531	thrombospondin-rel
14	59	100.0	574	2 A46283	sporozoite surface
15	59	100.0	2165	2 T21371	hypothetical prote
16	56	94.9	343	2 A29319	circumsporozoite p
17	56	94.9	351	1 OZ2QKU	circumsporozoite p
18	56	94.9	363	1 OZ2QAK	circumsporozoite p
19	56	94.9	367	2 A32068	circumsporozoite p
20	56	94.9	378	1 OZ2QAV	circumsporozoite p
21	56	94.9	386	2 A48571	circumsporozoite p
22	56	94.9	387	2 D41156	circumsporozoite p
23	56	94.9	387	2 C41156	circumsporozoite p
24	56	94.9	395	2 A41156	circumsporozoite p
25	55	93.2	805	2 T34212	hypothetical prote
26	55	93.2	807	2 A38152	F-spondin - rat
27	55	93.2	1074	2 JC5928	semaphorin F precu
28	55	93.2	1444	2 T18856	angiogenesis inhib
29	53	89.8	951	2 T00017	gene ADAMTS-1 prot

ALIGNMENTS

RESULT 1

OZ2QAB

circumsporozoite protein precursor - Plasmodium cynomolgi (strain Berok)
N:Alternate names: major sporozoite surface antigen
C:Species: Plasmodium cynomolgi
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
C:Accession: D26255
R:Galinski, M.R.; Arnot, D.E.; Cochrane, A.H.; Nussenzweig, R.S.; Ene
Cell 48, 311-319, 1987
A:Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
A:Reference number: A30889; MUID:87102878
A:Accession: D26255
A:Molecule type: DNA
A:Residues: 1-378 <GAL>
C:Comment: There are three distinct regions in the mature circumsporozoite protein, t
obc membrane-anchoring sequence.
C:Comment: There are 10 tandem copies of a 9-residue repeat (preceded by a 6-residue
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: sporozoite; surface antigen; tandem repeat
F:1-19/Domain: signal sequence #status predicted <Sig>
F:20-378/Product: circumsporozoite protein #status predicted <MAT>
F:97-192/Region: 9-residue repeats
F:193-268/Region: 16-residue repeats
F:303-356/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 59; DB 1; Length 378;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WSPCSVTCG 9

Db 313 WSPCSVTCG 321

RESULT 2

OZ2QAL

circumsporozoite protein precursor - Plasmodium cynomolgi (strain London)
N:Alternate names: major sporozoite surface antigen
C:Species: Plasmodium cynomolgi
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
C:Accession: A26255
R:Galinski, M.R.; Arnot, D.E.; Cochrane, A.H.; Nussenzweig, R.S.; Ene
Cell 48, 311-319, 1987
A:Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
A:Reference number: A30889; MUID:87102878
A:Accession: A26255
A:Molecule type: DNA
A:Residues: 1-378 <GAL>
C:Comment: There are three distinct regions in the mature circumsporozoite protein, t
obc membrane-anchoring sequence.
C:Comment: There are 19 tandem copies of a 6-residue repeat and 6 copies of an 11-res

C.Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C.Keywords: sporozoite; surface antigen; tandem repeat
 F.1-19/Domain: signal sequence #status predicted <SIG>
 F.20-378/Product: circumsporozoite protein #status predicted <MAT>
 F.98-211/Region: 6-residue repeats
 F.212-277/Region: 11-residue repeats
 F.303-356/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 100.0%; Score 59; DB 1; Length 378;
 Best Local Similarity 100.0%; Pred. No. 0.036;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WSPCSVTCG 9
 DB 313 WSPCSVTCG 321

RESULT 3
 A39756
 circumsporozoite protein - Plasmodium reichenowi
 C.Species: Plasmodium reichenowi
 C.Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 20-Aug-1999
 C.Accession: A39756
 C.Alternate names: major sporozoite surface antigen
 J. Biol. Chem. 266: 6686-6689, 1991
 A.Title: Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malaria
 A.Reference number: A39756; MUID: 91201303
 A.Accession: A39756
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-388 <IAL>
 A.Cross-references: GB:M60972; NID:AAA29561.1; PID:g160229
 C.Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 F.312-366/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 100.0%; Score 59; DB 2; Length 388;
 Best Local Similarity 100.0%; Pred. No. 0.037;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WSPCSVTCG 9
 DB 322 WSPCSVTCG 330

RESULT 4
 OZQAS
 circumsporozoite protein precursor - Plasmodium cynomolgi (strain Ceylon)
 N.Alternate names: major sporozoite surface antigen
 C.Species: Plasmodium cynomolgi
 C.Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
 C.Accession: C26255
 R.Galinski, M.R.; Arnot, D.E.; Cochrane, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Enea,
 Cell 48, 311-319, 1987
 A.Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
 A.Reference number: A90889; MUID: 87102878
 A.Accession: C26255
 A.Molecule type: DNA
 A.Residues: 1-398 <GAL>
 C.Comment: There are three distinct regions in the mature circumsporozoite protein, the
 obic membrane-anchoring sequence.
 C.Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C.Keywords: sporozoite; surface antigen; tandem repeat
 F.1-19/Domain: signal sequence #status predicted <SIG>
 F.97-240/Product: circumsporozoite protein #status predicted <MAT>
 F.97-240/Region: 9-residue repeats
 F.241-294/Region: 17-residue repeats
 F.323-376/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 100.0%; Score 59; DB 1; Length 398;

Best Local Similarity 100.0%; Pred. No. 0.038;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WSPCSVTCG 9
 DB 333 WSPCSVTCG 341

RESULT 5
 OZQAC
 circumsporozoite protein precursor - Plasmodium cynomolgi (strain Gombak)
 N.Alternate names: major sporozoite surface antigen
 C.Species: Plasmodium cynomolgi
 C.Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
 C.Accession: E26255
 R.Galinski, M.R.; Arnot, D.E.; Cochrane, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Ene
 Cell 48, 311-319, 1987
 A.Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
 A.Reference number: A90889; MUID: 87102878
 A.Accession: E26255
 A.Molecule type: DNA
 A.Residues: 1-401 <GAL>
 C.Comment: There are three distinct regions in the mature circumsporozoite protein, t
 obic membrane-anchoring sequence.
 C.Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C.Keywords: sporozoite; surface antigen; tandem repeat
 F.1-19/Domain: signal sequence #status predicted <SIG>
 F.20-401/Product: circumsporozoite protein #status predicted <MAT>
 F.98-278/Region: 11-residue repeats
 F.326-379/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 100.0%; Score 59; DB 1; Length 401;
 Best Local Similarity 100.0%; Pred. No. 0.038;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
 DB 336 WSPCSVTCG 344

RESULT 6
 S05428
 circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate NF54)
 C.Species: Plasmodium falciparum
 C.Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jun-2000
 C.Accession: S05428; A45527; I60657
 R.Campbell, J.R.
 Nucleic Acids Res. 17, 5854, 1989
 A.Title: DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate
 A.Reference number: S05428; MUID: 89345189
 A.Accession: S05428
 A.Status: translation not shown
 A.Molecule type: DNA
 A.Residues: 1-405 <GAL>
 R.Caspers, P.; Gentz, R.; Matile, H.; Pink, J.R.; Sinigaglia, F.
 Mol. Biochem. Parasitol. 35, 185-190, 1989
 A.Title: The circumsporozoite protein gene from NF54, a Plasmodium falciparum isolate
 A.Reference number: A45527; MUID: 89364998
 A.Accession: A45527
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-405 <CAS>
 A.Cross-references: GB:M2982; GB:J04650; NID:g160168; PID:AAA29527.1; PID:g160169
 R.Lockyer, M.J.; Marsh, K.; Newbold, C.;
 Mol. Biochem. Parasitol. 37, 27-30, 1989
 A.Title: Wild type and mutant Plasmodium falciparum show extensive polymorphism in T cell
 A.Reference number: A60657; MUID: 90114334
 A.Accession: I60657
 A.Status: preliminary; not compared with conceptual translation
 A.Molecule type: DNA

A:Residues: 319-336,354-373 <LOC>
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 F:329-383/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 100.0%; Score 59; DB 2; Length 405;
 Best Local Similarity 100.0%; Pred. No. 0.039;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTCG 9
 Db 339 WSPCSVTCG 347
 |||||

RESULT 7

OZZQAF
 circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate INTW22)
 C:Species: Plasmodium falciparum
 C:Date: 15-Nov-1984 #sequence_revision 15-Nov-1984 #text_change 09-Jun-2000
 C:Accession: A03388
 R:Dame, J.B.; Williams, J.L.; McCutchan, T.F.; Weber, J.L.; Wirtz, R.A.; Hochmeyer, W.T.
 Science 225, 593-599, 1984
 A:Title: Structure of the gene encoding the immunodominant surface antigen on the sporozoite
 A:Reference number: A03388; MUID:84250215
 A:Accession: A03388
 A:Molecule type: DNA
 A:Residues: 1-412 <DAM>
 A:Cross-references: GB:K02194; NID:gl60160; PIDN:AAA29524.1; PID:gl60161
 A:Experimental source: clone 7c8
 A:Comment: Residues 1-16 are the probable signal sequence.
 C:Comment: There are 41 copies of a 4-residue repeating unit in the middle domain of the protein.
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 F:336-390/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 100.0%; Score 59; DB 1; Length 412;
 Best Local Similarity 100.0%; Pred. No. 0.039;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTCG 9
 Db 346 WSPCSVTCG 354
 |||||

RESULT 8

OZZQAM
 circumsporozoite protein precursor - Plasmodium cynomolgi (strain Mulligan/NIH)
 N:Alternate names: major sporozoite surface antigen
 C:Species: Plasmodium cynomolgi
 C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
 C:Accession: B26255
 R:Galinski, M.R.; Arnot, D.E.; Cochrane, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Enea, Cell 48, 311-319, 1987
 A:Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
 A:Reference number: A90889; MUID:87102878
 A:Accession: B26255
 A:Molecule type: DNA
 A:Residues: 1-419 <GAL>
 C:Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.
 C:Comment: There are 53 tandem copies of a 4-residue repeat.
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C:Keywords: sporozoite; surface antigen; tandem repeat
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-419/Product: circumsporozoite protein #status predicted <MAT>
 F:99-310/Region: 4-residue repeats
 F:344-397/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 100.0%; Score 59; DB 1; Length 419;
 Best Local Similarity 100.0%; Pred. No. 0.04;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTCG 9
 Db 354 WSPCSVTCG 362
 |||||

RESULT 9

A54533
 circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain T4, The C:Species: Plasmodium falciparum
 C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
 C:Accession: A54533
 R:del Portillo, H.A.; Nussenzweig, R.S.; Enea, V. Mol. Biochem. Parasitol. 24, 289-294, 1987
 A:Title: Circumsporozoite gene of a Plasmodium falciparum strain from Thailand.
 A:Reference number: A54533; MUID:87315205
 A:Accession: A54533
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-424
 A:Cross-references: GB:M19752; NID:gl60216; PIDN:AAA29555.1; PID:gl60217
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 F:348-402/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 100.0%; Score 59; DB 2; Length 424;
 Best Local Similarity 100.0%; Pred. No. 0.04;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTCG 9
 Db 358 WSPCSVTCG 366
 |||||

RESULT 10

A54504
 circumsporozoite protein - Plasmodium malariae (strain Uganda 1-CDC)
 C:Species: Plasmodium malariae
 C:Date: 06-Oct-1994 #sequence_revision 04-Nov-1994 #text_change 20-Aug-1999
 C:Accession: A54504
 R:Lal, A.A.; de la Cruz, V.F.; Campbell, G.H.; Procell, P.M.; Collins, W.E.; McCutcha Mol. Biochem. Parasitol. 30, 291-294, 1988
 A:Title: Structure of the circumsporozoite gene of Plasmodium malariae.
 A:Reference number: A54504; MUID:89040027
 A:Accession: A54504
 A:Molecule type: DNA
 A:Residues: 1-429 <LAL>
 A:Cross-references: GB:J03992; NID:gl60220; PIDN:AAA29557.1; PID:gl60221
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C:Keywords: tandem repeat
 F:354-407/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 100.0%; Score 59; DB 2; Length 429;
 Best Local Similarity 100.0%; Pred. No. 0.04;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTCG 9
 Db 364 WSPCSVTCG 372
 |||||

RESULT 11

A54529
 circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain Wellcome) C:Species: Plasmodium falciparum
 C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
 C:Accession: A54529
 R:Lockyer, M.J.; Schwarz, R.T. Mol. Biochem. Parasitol. 22, 101-108, 1987
 A:Title: Strain variation in the circumsporozoite protein gene of Plasmodium falcipar A:Reference number: A54529; MUID:87115616
 A:Accession: A54529
 A:Status: Preliminary; not compared with conceptual translation

A:Molecule type: DNA
A:Residues: 1-442 <LOC>
A:Cross-references: GB:M15505; NID:g160214; PIDN:AAA29554.1; PID:g160215
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: tandem repeat
F:366-420/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 59; DB 2; Length 442;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
|||||
DB 376 WSPCSVTCG 384

RESULT 12
A60610
circumsporozoite protein precursor - Plasmodium brasilianum
C:Species: Plasmodium brasilianum
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 20-Aug-1999
C:Accession: A60610; A28615
R:di Giovanni, L.; Cochrane, A.H.; Enea, V.
Exp. Parasitol. 70, 373-381, 1990
A:Title: On the evolutionary history of the circumsporozoite protein in plasmodia.
A:Reference number: A60610; MUID:90214818
A:Accession: A60610
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-485 <DIA>
A:Experimental source: strain Colombian
R:Lal, A.A.; de la Cruz, V.F.; Collins, W.E.; Campbell, G.H.; Procell, P.M.; McCutchan, J. Biol. Chem. 263, 5495-5498, 1988
A:Title: Circumsporozoite protein gene from Plasmodium brasilianum. Animal reservoirs for malaria.
A:Reference number: A28615; MUID:88186854
A:Accession: A28615
A:Molecule type: DNA
A:Residues: 93-485 <LAJ>
A:Cross-references: GB:J03203; NID:g160212; PIDN:AAA29553.1; PID:g160213
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: sporozoite; surface antigen; tandem repeat
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-485/Product: circumsporozoite protein #status predicted <MAT>
F:114-369/Region: 4-residue repeats
F:410-463/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 59; DB 2; Length 485;
Best Local Similarity 100.0%; Pred. No. 0.044;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
|||||
DB 420 WSPCSVTCG 428

RESULT 13
S04531
thrombospondin-related protein TRAP - malaria parasite (Plasmodium falciparum)
N:Alternate names: thrombospondin-related anonymous protein
C:Species: Plasmodium falciparum
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jun-2000
C:Accession: S04531
R:Robson, K.J.H.; Hall, J.R.S.; Jennings, M.W.; Harris, T.J.R.; Marsh, K.; Newbold, C.I. Nature 335, 79-82, 1988
A:Title: A highly conserved amino-acid sequence in thrombospondin, properdin and in protein 35.
A:Reference number: S04531; MUID:88318952
A:Accession: S04531
A:Molecule type: DNA
A:Residues: 1-559 <ROB>
A:Cross-references: EMBL:X13022; NID:g9977; PID:g9978
C:Superfamily: thrombospondin type 1 repeat homology

F:240-287/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 59; DB 2; Length 559;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
|||||
DB 250 WSPCSVTCG 258

RESULT 14
A46283
sporozoite surface protein 2 - malaria parasite (Plasmodium falciparum) (strain NF54)
N:Alternate names: thrombospondin-related anonymous protein (TRAP)
C:Species: Plasmodium falciparum
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jun-2000
C:Accession: A46283
R:Rogers, W.O.; Malik, A.; Mellouk, S.; Nakamura, K.; Rogers, M.D.; Szarfman, A.; Gor Proc. Natl. Acad. Sci. U.S.A. 89, 9176-9180, 1992
A:Title: Characterization of Plasmodium falciparum sporozoite surface protein 2.
A:Reference number: A46283; MUID:93028427
A:Accession: A46283
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-574 <ROG>
A:Cross-references: GB:M94013; NID:g160690; PID:g160691
C:Superfamily: thrombospondin type 1 repeat homology
F:240-287/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 59; DB 2; Length 574;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
|||||
DB 250 WSPCSVTCG 258

RESULT 15
T21371
hypothetical protein F25H8.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T21371; T24896
R:Gajadaty, S.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19413
A:Accession: T21371
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2165 <WIL>
A:Cross-references: EMBL:Z69360; PIDN:CAA93287.1; GSPDB:GN00022; CESP:F25H8.3
A:Experimental source: clone F25H8
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19949
A:Accession: T24896
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2165 <WIL>
A:Cross-references: EMBL:Z69361; PIDN:CAA93288.1; GSPDB:GN00022; CESP:F25H8.3
A:Experimental source: clone T13H10
C:Genetics:
A:Gene: CESP:F25H8.3
A:Map position: 4
A:introns: 31/1; 52/1; 135/2; 193/3; 216/1; 266/1; 495/2; 547/3; 584/3; 634/2; 744/1.

Query Match 100.0%; Score 59; DB 2; Length 2165;
Best Local Similarity 100.0%; Pred. No. 0.14;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 WSPCSVTCG 9
| | | | | | | | | |
Db 1304 WSPCSVTCG 1312

Search completed: January 29, 2002, 10:26:37
Job time: 648 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:13:39 ; Search time 80.65 seconds
(without alignments)
4.092 Million cell updates/sec

Title: US-09-763-397a-8
Perfect score: 59
Sequence: 1 WSPCSVTCG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 segs, 36664827 residues
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed;
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	59	100.0	378	1	CSP_PLACB
2	59	100.0	388	1	CSP_PLARE
3	59	100.0	393	1	CSP_PLABR
4	59	100.0	397	1	CSP_PLAFO
5	59	100.0	398	1	CSP_PLACC
6	59	100.0	412	1	CSP_PLAFA
7	59	100.0	419	1	CSP_PLACM
8	59	100.0	424	1	CSP_PLAFT
9	59	100.0	429	1	CSP_PLAMA
10	59	100.0	442	1	CSP_PLAFW
11	59	100.0	559	1	TRAP_PLAFA
12	56	94.9	343	1	CSP_PLAVS
13	56	94.9	351	1	CSP_PLAKU
14	56	94.9	363	1	CSP_PLAKH
15	56	94.9	378	1	CSP_PLACL
16	56	94.9	378	1	CSP_PLAVB
17	56	94.9	386	1	CSP_PLACG
18	56	94.9	401	1	CSP_PLACI
19	56	94.9	1173	1	TSP1_XENLA
20	55	93.2	807	1	FSP0_RAT
21	55	93.2	1074	1	SMSA_HUMAN
22	55	93.2	1077	1	SMSA_MOUSE
23	53	89.8	967	1	ATS1_RAT
24	53	89.8	968	1	ATS1_MOUSE
25	52	88.1	1629	1	ATS9_HUMAN
26	51	86.4	367	1	CSP_PLAYO
27	51	86.4	803	1	FSP0_XENLA
28	51	86.4	867	1	FSP0_BOVIN
29	51	86.4	1170	1	TSP1_BOVIN
30	51	86.4	1170	1	TSP1_HUMAN
31	51	86.4	1170	1	TSP1_MOUSE
32	51	86.4	1170	1	TSP2_BOVIN
33	51	86.4	1172	1	TSP2_HUMAN

34	51	86.4	1172	1	TSP2_MOUSE	O03350	mus musculus
35	51	86.4	1178	1	TSP2_CHICK	P35440	gallus gall
36	51	86.4	1201	1	ATS3_HUMAN	O15072	homo sapien
37	51	86.4	1205	1	ATS2_BOVIN	P79331	b adam-ts 2
38	51	86.4	1211	1	ATS2_HUMAN	O95450	h adam-ts 2
39	51	86.4	1522	1	BAI3_HUMAN	O60242	homo sapien
40	50	84.7	469	1	PROP_HUMAN	P27918	homo sapien
41	49	83.1	437	1	PROP_MOUSE	P11680	homo sapien
42	48	81.4	339	1	CSP_PLABE	P06915	plasmodium
43	48	81.4	347	1	CSP_PLABA	P23093	plasmodium
44	48	81.4	470	1	PROP_CAVPO	O64181	cavia porce
45	47	79.7	1572	1	BAI2_HUMAN	O60241	homo sapien

ALIGNMENTS

```
RESULT 1
CSP_PLACB
ID CSP_PLACB STANDARD; PRT; 378 AA.
AC P08672;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium cynomolgi (strain Berok).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_Taxid=5828;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87102878; PubMed=3802196;
RA Galinski M.R., Arnot D.E., Cochran A.H., Barnwell J.W.,
RA Nussenzweig R.S., Enea V.;
RT "The circumsporoite gene of the Plasmodium cynomolgi complex.";
RL Cell 48:311-319(1987).
CC -!- DOMAIN: THERE ARE 10 TANDEM COPIES OF A 9-RESIDUE REPEAT (PRECEDDED
CC BY A 6-RESIDUE INCOMPLETE REPEAT) AND 3 TANDEM COPIES OF A 16-
CC RESIDUE REPEAT (FOLLOWED BY 3 SHORTER, INCOMPLETE COPIES).
CC -!- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement. See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch.
-----
EMBL: M15104; AAA29532.1;
PIR: D26255; OZ20AB.
InterPro: IPR003067; Crcmsprzoite.
InterPro: IPR000884; TSP1.
Pfam: PF00090; tsp-1; 1.
PRINTS: PR01303; CRCMSPRZOITE.
SMART: SM00209; TSP1; 1.
Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19
FT CHAIN 20 378
FT DOMAIN 97 192
FT DOMAIN 193 240
FT REPEAT 241 251
FT REPEAT 252 260
FT REPEAT 261 268
FT REPEAT 261 268
SQ SEQUENCE 378 AA; 36286 MW; 779BA081C140793F CRC64;
```

Query Match 100.0%; Score 59; DB 1; Length 378;
Best Local Similarity 100.0%; Pred. No. 0.009;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
| | | | | | | | | |
DB 313 WSPCSVTCG 321

RESULT 2

CSP_PLAIRE STANDARD; PRT; 388 AA.
ID P26694;
DT 01-AUG-1992 (Rel. 23, Created)
DE 01-AUG-1992 (Rel. 23, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5854;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91201303; PubMed=2016283;
RA Lal A.A., Goldman I.F.;
RT "Circumsporoite protein gene from Plasmodium reichenowi, a
RT chimpanzee malaria parasite evolutionarily related to the human
RT malaria parasite Plasmodium falciparum.";
RL J. Biol. Chem. 266:6686-6689(1991).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M60972; AAA29561.1; -
CC PIR; A39756; A39756.
CC InterPro; IPR003067; Crcmsprzoite.
CC InterPro; IPR000884; TSP1.
CC Pfam; PF00090; tsp_1; 1.
CC PRINTS; PR01303; CRCMSPRZOITE.
CC SMART; SM00209; TSP1; 1.
CC Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 CIRCUMSPOROZOITE PROTEIN.
FT CHAIN 17 388
FT DOMAIN 120 267 62 X 4 AA TANDEM REPEATS OF N-A-N-P.
SQ SEQUENCE 388 AA; 42245 MW; C031EFFBE2E35604 CRC64;

Query Match 100.0%; Score 59; DB 1; Length 388;
Best Local Similarity 100.0%; Pred. No. 0.0092;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
| | | | | | | | | |
DB 322 WSPCSVTCG 330

RESULT 3

CSP_PLABR STANDARD; PRT; 393 AA.
ID P14593;
AC P14593;
DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (CS) (FRAGMENT).
OS Plasmodium brasilianum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5824;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88186854; PubMed=3128542;
RA Lal A.A., la Cruz V.F., Collins W.E., Campbell G.H., Procell P.M.,
RA McCutchan T.F.;
RT "Circumsporoite protein gene from Plasmodium brasilianum. Animal
RT reservoirs for human malaria parasites?";
RL J. Biol. Chem. 263:5495-5498(1988).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J03203; AAA29553.1; -
CC PIR; A28615; A28615.
CC HSP; P04002; IWFA.
CC InterPro; IPR000884; TSP1.
CC Pfam; PF00090; tsp_1; 1.
CC SMART; SM00209; TSP1; 1.
CC PROSITE; PS00092; TSP1; 1.
CC Malaria; Repeat; Sporozoite.
FT NON_TER 1 1
FT DOMAIN 16 278 TANDEM REPEATS OF G-N-A-A.
SQ SEQUENCE 393 AA; 35372 MW; B32944419BC600AA CRC64;

Query Match 100.0%; Score 59; DB 1; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.0093;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
| | | | | | | | | |
DB 328 WSPCSVTCG 336

RESULT 4

CSP_PLAFO STANDARD; PRT; 397 AA.
ID P19597; Q25798;
AC P19597; Q25798;
DT 01-FEB-1991 (Rel. 17, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium falciparum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5843;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89345189; PubMed=2668895;
RA Campbell J.R.;
RT "DNA sequence of the gene encoding a Plasmodium falciparum malaria
RT candidate vaccine antigen.";
RL Nucleic Acids Res. 17:5854-5854(1989).
RN [2]
RP REVISIONS.

RA Campbell J.R.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92155298; PubMed=1346766;
RA Davis J.R., Cortese J.F., Herrington D.A., Murphy J.R., Clyde D.F.,
RA Thomas A.W., Bagar S., Cochran M.A., Thanassi J., Levine M.M.,
RA Hackett C.S.;
RT "Plasmodium falciparum: in vitro characterization and human
RT infectivity of a cloned line";
RL Exp. Parasitol. 74:159-168(1992).
[4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89364998; PubMed=2671723;
RA Caspers P., Gentz R., Matile H., Pink J.R., Sinigaglia F.;
RT "The circumsporozoite protein gene from NF54, a Plasmodium falciparum
RT isolate used in malaria vaccine trials";
RL Mol. Biochem. Parasitol. 35:185-190(1989).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15363; CAA33421.1; -;
DR EMBL; M83886; AAA29521.1; -;
DR EMBL; M22982; AAA29527.1; -;
DR PIR; S05428; S05428.
DR PIR; A45527; A45527.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSPI.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPI; 1.
DR Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 397 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 101 272 43 X 4 AA TANDEM REPEATS OF N-A-N-P.
FT CONFLICT 194 194 A -> ANPNANPNA (IN REF. 4).
SQ SEQUENCE 397 AA; 42646 MW; 9E81146F59EBCEA3 CRC64;

Query Match 100.0%; Score 59; DB 1; Length 397;
Best Local Similarity 100.0%; Pred. No. 0.0093;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTGC 9
Db 331 WSPCSVTGC 339
|||||||

RESULT 5
CSP_PLACC STANDARD; PRT; 398 AA.
AC P08673;
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium cynomolgi (strain Ceylon).
CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
CC NCBI_TaxID=5829;

SEQUENCE FROM N.A.
MEDLINE=84250215; PubMed=6204383;
Dane J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
Hochmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
Sanders G.S., Reddy E.P., Diggs C.L., Miller L.H.;
"Structure of the gene encoding the immunodominant surface antigen on
the sporozoite of the human malaria parasite Plasmodium falciparum";
Science 225:593-599(1984).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR

RP SEQUENCE FROM N.A.
RX MEDLINE=87102878; PubMed=3802196;
RA Galinski M.R., Arnot D.E., Cochran A.H., Barnwell J.W.,
RA Nussenzweig R.S., Enea V.;
RT "The circumsporozoite gene of the Plasmodium cynomolgi complex";
RL Cell 48:311-319(1987).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15103; AAA29533.1; -;
DR PIR; C26255; OZQOAS.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSPI.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPI; 1.
DR Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19
FT CHAIN 20 398 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 97 249 17 X 9 AA TANDEM REPEATS OF A-G-N-N-A-A-
FT A-G-E.
SQ SEQUENCE 398 AA; 37718 MW; 6DFA2E8A62ED05BF CRC64;

Query Match 100.0%; Score 59; DB 1; Length 398;
Best Local Similarity 100.0%; Pred. No. 0.0094;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTGC 9
Db 333 WSPCSVTGC 341
|||||||

RESULT 6
CSP_PLAFA STANDARD; PRT; 412 AA.
AC P02893;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium falciparum.
CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
CC NCBI_TaxID=5833;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84250215; PubMed=6204383;
Dane J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
Hochmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
Sanders G.S., Reddy E.P., Diggs C.L., Miller L.H.;
"Structure of the gene encoding the immunodominant surface antigen on
the sporozoite of the human malaria parasite Plasmodium falciparum";
Science 225:593-599(1984).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR

CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: K02194; AAA29524.1; -.
CC PIR: A03388; OZQZAF.
CC InterPro: IPR003067; Crcmsprzoite.
CC Pfam: PF00090; tsp_1; 1.
CC PRINTS: PR01303; Crcmsprzoite.
CC SMART: SM00209; TSP1; 1.
CC Malaria: Sporozoite; Repeat; Signal.
CC SIGNAL 1 16 PROBABLE.
CC FT CHAIN 17 412 CIRCUMSPOROZOITE PROTEIN.
CC FT DOMAIN 123 290 41 X 4 AA TANDEM REPEATS OF P-N-A-N.
CC SEQUENCE 412 AA; 44420 MW; 1EEED3DE0965F8 CRC64;
CC
CC Query Match 100.0%; Score 59; DB 1; Length 412;
CC Best Local Similarity 100.0%; Pred. No. 0.0096;
CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 WSPCSVTGC 9
CC |||||
CC DB 346 WSPCSVTGC 354
CC
CC RESULT 7
CC CSP_PLACM STANDARD; PRT; 419 AA.
CC AC P08676;
CC DT 01-JAN-1988 (Rel. 06, Created)
CC DT 01-FEB-1996 (Rel. 33, Last sequence update)
CC DT 20-AUG-2001 (Rel. 40, Last annotation update)
CC DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
CC OS Plasmodium cynomolgi (strain Mulligan/NIH).
CC OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
CC OX NCBI_TaxID=5832;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=87102878; PubMed=3802196;
CC RA Galinski M.R., Arnot D.E., Cochran A.H., Barnwell J.W.,
CC RA Nussenzweig R.S., Enea V.;
CC RT "The circumsporoite gene of the plasmodium cynomolgi complex.";
CC RL Cell 48:311-319(1987).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M15102; AAA29539.1; -.
CC PIR: B26255; OZQZAM.
CC InterPro: IPR003067; Crcmsprzoite.
CC

DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; Crcmsprzoite.
DR SMART: SM00209; TSP1; 1.
DR Malaria: Sporozoite; Repeat; Signal.
DR SIGNAL 1 19 PROBABLE.
DR FT CHAIN 20 419 CIRCUMSPOROZOITE PROTEIN.
DR FT DOMAIN 99 314 54 X 4 AA TANDEM REPEATS OF N-A-[DG]-G.
DR SEQUENCE 419 AA; 38924 MW; 8F46CDD8A1B4EFF4 CRC64;
DR
DR Query Match 100.0%; Score 59; DB 1; Length 419;
DR Best Local Similarity 100.0%; Pred. No. 0.0098;
DR Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DR
DR QY 1 WSPCSVTGC 9
DR |||||
DR DB 354 WSPCSVTGC 362
DR
DR RESULT 8
DR CSP_PLAFT STANDARD; PRT; 424 AA.
DR AC P13814;
DR DT 01-JAN-1990 (Rel. 13, Created)
DR DT 01-JAN-1990 (Rel. 13, Last sequence update)
DR DT 20-AUG-2001 (Rel. 40, Last annotation update)
DR DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
DR OS Plasmodium falciparum (isolate t4 / Thailand).
DR OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
DR OX NCBI_TaxID=5846;
DR RN [1]
DR RP SEQUENCE FROM N.A.
DR RX MEDLINE=87315205; PubMed=3306373;
DR RA del Portillo H.A., Nussenzweig R.S., Enea V.;
DR RT "Circumsporoite gene of a Plasmodium falciparum strain from
DR Thailand.";
DR RL Mol. Biochem. Parasitol. 24:289-294(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M19752; AAA29555.1; -.
CC PIR: A54533; A54533.
CC InterPro: IPR003067; Crcmsprzoite.
CC Pfam: PF00090; tsp_1; 1.
CC PRINTS: PR01303; Crcmsprzoite.
CC SMART: SM00209; TSP1; 1.
CC Malaria: Sporozoite; Repeat; Signal.
CC SIGNAL 1 16 PROBABLE.
CC FT CHAIN 17 424 CIRCUMSPOROZOITE PROTEIN.
CC FT DOMAIN 123 300 45 X 4 AA TANDEM REPEATS OF N-A-N-P.
CC SEQUENCE 424 AA; 45610 MW; 710AB14238786CD9 CRC64;
CC
CC Query Match 100.0%; Score 59; DB 1; Length 424;
CC Best Local Similarity 100.0%; Pred. No. 0.0099;
CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 WSPCSVTCG 9
DB 358 WSPCSVTCG 366

RESULT 9
CSP_PLAMA
ID CSP_PLAMA STANDARD; PRT; 429 AA.
AC P13815;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5858;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89040027; PubMed=3054537;
RA Lal A.A., la Cruz V.F., Campbell G.H., Procell P.M., Collins W.E.,
RA McCutchan T.F.;
RL Mol. Biochem. Parasitol. 30:291-294 (1988).
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03992; AAA29557.1; -
DR PIR; A54504; A54504.
DR HSP; P04002; LWFA.
DR InterPro; IPR003067; Crcmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS00092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 15 PROBABLE.
FT CHAIN 16 429 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 110 314 4 AA TANDEM REPEATS.
FT SEQUENCE 429 AA; 41596 MW; 3629D641D1C0BB7E CRC64;

Query Match 100.0%; Score 59; DB 1; Length 429;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
DB 364 WSPCSVTCG 372

RESULT 10
CSP_PLAFW
ID CSP_PLAFW STANDARD; PRT; 442 AA.
AC P08307;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).

```

```

OS Plasmodium falciparum (isolate Wellcome).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5848;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87115616; PubMed=3543671;
RA Lockyer M.J., Schwarz R.T.;
RT "Strain variation in the circumsporoite protein gene of Plasmodium
RT falciparum.";
RL Mol. Biochem. Parasitol. 22:101-108 (1987).
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15505; AAA29554.1; -
DR PIR; A54529; A54529.
DR InterPro; IPR003067; Crcmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 442 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 130 320 47 X 4 AA TANDEM REPEATS OF N-A-N-P.
FT SEQUENCE 442 AA; 47402 MW; BD57A9A152B85E03 CRC64;

Query Match 100.0%; Score 59; DB 1; Length 442;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
DB 376 WSPCSVTCG 384

RESULT 11
TRAP_PLAFA
ID TRAP_PLAFA STANDARD; PRT; 559 AA.
AC P16893;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE THROMBOSPONDIN RELATED ANONYMOUS PROTEIN PRECURSOR.
GN TRAP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86318952; PubMed=3045563;
RA Robson K.J.H., Hall J.R.S., Jennings M.W., Harris T.J.R.,
RA Marsh K., Newbold C.I., Tate V.E., Weatherall D.J.;
RT "A highly conserved amino-acid sequence in thrombospondin, properdin
RT and in proteins from sporozoites and blood stages of a human malaria
RT parasite.";
RL Nature 335:79-82 (1988).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED DURING ERYTHROCYTIC STAGE OF
CC LIFE CYCLE.

```

```
CC -!- SIMILARITY: IT IS POSSIBLE THAT TRAP IS THE P.FALCIPARUM HOMOLOG
CC OF THE 66 KDA PROTEIN FROM P.KNOWLESI.
CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 VWFA DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X13022; CAA31440.1; -
DR PIR; S04531;
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS50234; VWFA; 1.
DR Antigen; Malaria; Glycoprotein; Cell adhesion; Signal; Transmembrane.
KW SIGNAL; 1; 2;
FT CHAIN; ? 559
FT DOMAIN; 48 234
FT DOMAIN; 241 288
FT SITE; 76 78
FT SITE; 307 309
FT CARBOHYD; 132 132
FT CARBOHYD; 310 310
FT CARBOHYD; 460 460
FT CARBOHYD; 468 468
SQ SEQUENCE 559 AA; 63300 MW; 02A20164D198C37A CRC64;

Query Match 100.0%; Score 59; DB 1; Length 559;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTGC 9
| | | | | | | |
DB 250 WSPCSVTGC 258

RESULT 12
CSP_PLAYS STANDARD; PRT; 343 AA.
AC P13826;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (CS) (FRAGMENT).
OS Plasmodium vivax (strain Salvador I)
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5856;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86070222; PubMed=2416057;
RA McCutchan T.F., Lal A.A., de la Cruz V.F., Miller L.H., Maloy W.L.,
RA Charoenwit Y., Beaudoin R.L., Guerry P., Wistar R. Jr., Hoffman S.L.,
RA Hockmeyer W.T., Collins W.E., Wirth D.;
RT "Sequence of the immunodominant epitope for the surface protein on
RT sporozoites of Plasmodium vivax."
RL Science 230:1381-1383(1985).
RP SEQUENCE FROM N.A.
RX MEDLINE=87194878; PubMed=2437120;
RA de la Cruz V.F., Lal A.A., Welsh J.A., McCutchan T.F.;
RA "Evolution of the immunodominant domain of the circumsporozoite
RT protein gene from Plasmodium vivax. Implications for vaccines.";
RT J. Biol. Chem. 262:6468-6467(1987).
```

```
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Sporozoite; Malaria; Repeat.
FT NON-TER 1
FT DOMAIN 63 243
FT SEQUENCE 343 AA; 34155 MW; 308EFD5B8C15DFC3 CRC64;
G-Q-P.
SQ SEQUENCE 343 AA; 34155 MW; 308EFD5B8C15DFC3 CRC64;

Query Match 94.9%; Score 56; DB 1; Length 343;
Best Local Similarity 88.9%; Pred. No. 0.025;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTGC 9
| | | | | | | |
DB 278 WTPCSVTGC 286

RESULT 13
CSP_PLAKU STANDARD; PRT; 351 AA.
AC P04922;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium knowlesi (strain nuri).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5852;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85272582; PubMed=4023712;
RA Sharma S., Svec P., Mitchell G.H., Godson G.N.;
RT "Diversity of circumsporozoite antigen genes from two strains of the
RT malarial parasite Plasmodium knowlesi."
RL Science 229:779-782(1985).
CC -!- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M11031; AAA29540.1; -
DR PIR; A26253; OZZQKU.
DR InterPro; IPR003067; Circmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 351
FT CHAIN 20 351
CIRCUMSPOROZOITE PROTEIN.
```

FT DOMAIN 101 235 15 X 9 AA TANDEM REPEATS OF A-A-G-A-G-G-
E-Q-P.
FT SEQUENCE 351 AA; 34782 MW; A85E87A152E6485B CRC64;

Query Match 94.9%; Score 56; DB 1; Length 351;
Best Local Similarity 88.9%; Pred. No. 0.026;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTGC 9
I:|||||
Db 286 WTPCSVTGC 294

RESULT 14
CSP_PLAKH STANDARD; PRT; 363 AA.
AC P02894;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium knowlesi (strain H).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5851;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84026486; PubMed=63113209;
RA Ozaki L.S., Svec P., Nussenzweig R.S., Nussenzweig V., Godson G.N.;
RT "Structure of the Plasmodium knowlesi gene coding for the
RT "circumsporoite protein.";
RL Cell 34:815-822(1983).
RN [2]
RP SEQUENCE OF 84-258 FROM N.A.
RX MEDLINE=63297689; PubMed=6193427;
RA Godson G.N., Ellis J., Svec P., Schlesinger D.H., Nussenzweig V.;
RT "Identification and chemical synthesis of a tandemly repeated
RT immunogenic region of Plasmodium knowlesi circumsporoite protein.";
RL Nature 305:29-33(1983).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; K00822; AAA19699.1; -
DR EMBL; K00772; AAA29556.1; -
PIR; A03389; OZZQAK.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
PFam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19
FT CHAIN 20 363 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 98 241 12 X 12 AA TANDEM REPEATS OF N-A-G-Q-P-Q-
A-Q-G-D-G-A.
SQ SEQUENCE 363 AA; 36793 MW; 574DF4BD320A7955 CRC64;

Query Match 94.9%; Score 56; DB 1; Length 363;
Best Local Similarity 88.9%; Pred. No. 0.027;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTGC 9
I:|||||
Db 298 WTPCSVTGC 306

RESULT 15
CSP_PLACL STANDARD; PRT; 378 AA.
AC P08675;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium cynomolgi (strain London).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5831;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87102878; PubMed=38021196;
RA Galinski M.R., Arnot D.E., Cochrane A.H., Barnwell J.W.,
RA Nussenzweig R.S., Enea V.;
RT "The circumsporoite gene of the Plasmodium cynomolgi complex.";
RL Cell 48:311-319(1987).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M15101; AAA29537.1; -
DR PIR; A26255; OZZQAL.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
PFam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 378 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 98 211 18 X 6 AA TANDEM REPEATS OF D-G-A-R-A-
[EA]
FT DOMAIN 212 277 6 X 11 AA TANDEM REPEATS OF G-N-Q-A-G-G-
Q-A-G-A-G.
SQ SEQUENCE 378 AA; 37462 MW; 8295A913C36420C5 CRC64;

Query Match 94.9%; Score 56; DB 1; Length 378;
Best Local Similarity 88.9%; Pred. No. 0.027;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTGC 9
I:|||||
Db 313 WSPCIVTGC 321

Search completed: January 29, 2002, 11:13:39
Job time: 815 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: January 29, 2002, 11:12:10 ; Search time 285.36 Seconds
(without alignments)
4.613 Million cell updates/sec

Title: US-09-763-397A-8
Perfect score: 59
Sequence: 1 WSPCSVTGC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	73	5 Q25833	Q25833 plasmodium
2	59	100.0	79	5 Q9U0Q2	Q9U0Q2 plasmodium
3	59	100.0	80	5 Q9U0Q4	Q9U0Q4 plasmodium
4	59	100.0	80	5 Q9U0Q3	Q9U0Q3 plasmodium
5	59	100.0	80	5 Q9U0Q1	Q9U0Q1 plasmodium
6	59	100.0	80	5 Q9U0Q0	Q9U0Q0 plasmodium
7	59	100.0	80	5 Q9U0P9	Q9U0P9 plasmodium
8	59	100.0	80	5 Q9U0P8	Q9U0P8 plasmodium
9	59	100.0	80	5 Q9U0P7	Q9U0P7 plasmodium
10	59	100.0	80	5 Q9U0P6	Q9U0P6 plasmodium
11	59	100.0	80	5 Q9U0P5	Q9U0P5 plasmodium
12	59	100.0	80	5 Q9U0P4	Q9U0P4 plasmodium
13	59	100.0	80	5 Q9U0P3	Q9U0P3 plasmodium
14	59	100.0	80	5 Q9U0P2	Q9U0P2 plasmodium
15	59	100.0	80	5 Q9U0P1	Q9U0P1 plasmodium
16	59	100.0	80	5 Q9U0P0	Q9U0P0 plasmodium
17	59	100.0	80	5 Q9U0W7	Q9U0W7 plasmodium
18	59	100.0	80	5 Q9U0W8	Q9U0W8 plasmodium
19	59	100.0	80	5 Q9U0W6	Q9U0W6 plasmodium
					Q9U0W1 plasmodium

20	59	100.0	80	5 Q9TW00	Q9TW00 plasmodium
21	59	100.0	80	5 Q9TVQ0	Q9TVQ0 plasmodium
22	59	100.0	80	5 Q9TVQ9	Q9TVQ9 plasmodium
23	59	100.0	80	5 Q9TVN9	Q9TVN9 plasmodium
24	59	100.0	106	5 Q9U6P2	Q9U6P2 plasmodium
25	59	100.0	106	5 Q9U6P1	Q9U6P1 plasmodium
26	59	100.0	106	5 Q9U6P0	Q9U6P0 plasmodium
27	59	100.0	115	5 Q25835	Q25835 plasmodium
28	59	100.0	115	5 Q25836	Q25836 plasmodium
29	59	100.0	115	5 Q25837	Q25837 plasmodium
30	59	100.0	115	5 Q25839	Q25839 plasmodium
31	59	100.0	115	5 Q9U934	Q9U934 plasmodium
32	59	100.0	115	5 Q9U937	Q9U937 plasmodium
33	59	100.0	117	5 Q25794	Q25794 plasmodium
34	59	100.0	117	5 Q25795	Q25795 plasmodium
35	59	100.0	117	5 Q25796	Q25796 plasmodium
36	59	100.0	117	5 Q25797	Q25797 plasmodium
37	59	100.0	168	5 Q9G223	Q9G223 cryptospori
38	59	100.0	168	5 Q9G222	Q9G222 cryptospori
39	59	100.0	168	5 Q9G221	Q9G221 cryptospori
40	59	100.0	383	5 Q9GPN1	Q9GPN1 plasmodium
41	59	100.0	393	5 Q9G255	Q9G255 plasmodium
42	59	100.0	408	5 Q25729	Q25729 plasmodium
43	59	100.0	416	5 Q25829	Q25829 plasmodium
44	59	100.0	420	5 Q25831	Q25831 plasmodium
45	59	100.0	420	5 Q25838	Q25838 plasmodium

ALIGNMENTS

```

RESULT 1
Q25833 PRELIMINARY; PRT; 73 AA.
ID AC Q25833
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CSP.
OS Plasmodium falciparum (isolate le5).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5840;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LE5;
RX MEDLINE=89009757; PubMed=2459211;
RA la Cruz V.F., Maloy W.L., Miller L.H., Lal A.A., Good M.F.,
RA McCutchan T.F.;
RT "Lack of cross-reactivity between variant T cell determinants from
RT malaria circumsporozoite protein.";
RL J. Immunol. 141:2456-2460(1988).
DR EMBL: W23185; AAA57043.1;
DR InterPro: IPR000884; TSPL;
DR InterPro: IPR003067; Circmsprzoite.
DR Pfam: PF00090; tspl_1;
DR PRINTS: PRO1303; Circmsprzoite.
DR SMART: SM00209; TSPL; 1.
FT NON_TER 1
FT NON_TER 73
SQ SEQUENCE 73 AA: 8233 MW: 53524145A58E52E8 CRC64:

```

Query Match 100.0%; Score 59; DB 5; Length 73;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WSPCSVTGC 9

Db 26 WSPCSVTGC 34

RESULT 2

Q9U0Q2
ID Q9U0Q2 PRELIMINARY; PRT; 79 AA.
AC Q9U0Q2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D51;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269945; CAB64243.1; -
DR InterPro: IPR000884; TSPI.
DR InterPro: IPR003067; Crmsprzoite.
DR Pfam: PF00090; tsp.1; 1.
DR PRINTS: PR01303; CRMSPRZOITE.
DR SMART: SM00209; TSPI; 1.
FT NON_TER 1 79
FT NON_TER 79 79
SQ SEQUENCE 79 AA; 8893 MW; 56FBFA76D859B416 CRC64;

Query Match 100.0%; Score 59; DB 5; Length 79;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
| | | | | | | | | |
Db 28 WSPCSVTCG 36

RESULT 3
Q9U0Q4
ID Q9U0Q4 PRELIMINARY; PRT; 80 AA.
AC Q9U0Q4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BL;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269941; CAB64167.1; -
DR InterPro: IPR000884; TSPI.
DR InterPro: IPR003067; Crmsprzoite.
DR Pfam: PF00090; tsp.1; 1.
DR PRINTS: PR01303; CRMSPRZOITE.
DR SMART: SM00209; TSPI; 1.
FT NON_TER 1 80
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9102 MW; A3283B70CEE50FDE CRC64;

Query Match 100.0%; Score 59; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
| | | | | | | | | |
Db 28 WSPCSVTCG 36

RESULT 4
Q9U0Q3
ID Q9U0Q3 PRELIMINARY; PRT; 80 AA.
AC Q9U0Q3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D25;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269942; CAB64168.1; -
DR InterPro: IPR000884; TSPI.
DR InterPro: IPR003067; Crmsprzoite.
DR Pfam: PF00090; tsp.1; 1.
DR PRINTS: PR01303; CRMSPRZOITE.
DR SMART: SM00209; TSPI; 1.
FT NON_TER 1 80
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9014 MW; 077C7E319AFE13B3 CRC64;

Query Match 100.0%; Score 59; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
| | | | | | | | | |
Db 28 WSPCSVTCG 36

RESULT 5
Q9U0Q1
ID Q9U0Q1 PRELIMINARY; PRT; 80 AA.
AC Q9U0Q1;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum (isolate K1 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269946; CAB64171.1; -
DR InterPro: IPR000884; TSPI.
DR InterPro: IPR003067; Crmsprzoite.
DR Pfam: PF00090; tsp.1; 1.
DR PRINTS: PR01303; CRMSPRZOITE.
DR PROSITE: PS00092; TSPI; 1.
DR SMART: SM00209; TSPI; 1.

```
FT NON_TER 1 1
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9137 MW; 4204EEG68F0B1434 CRC64;

Query Match 100.0%; Score 59; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTCG 9
Db 28 WSPCSVTCG 36

RESULT 6
Q9U0Q0 PRELIMINARY; PRT; 80 AA.
AC Q9U0Q0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
CN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RC STRAIN=D28;
RP SEQUENCE FROM N.A.
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ269948; CAB64173.1; -.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1 1
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9046 MW; BA7689D18F031C3E CRC64;

Query Match 100.0%; Score 59; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTCG 9
Db 28 WSPCSVTCG 36

RESULT 7
Q9U0P9 PRELIMINARY; PRT; 80 AA.
AC Q9U0P9;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
CN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RC STRAIN=Li;
RP SEQUENCE FROM N.A.
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ269951; CAB64176.1; -.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1 1
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9032 MW; ADED6F0E266AD98E CRC64;
```

```
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ269951; CAB64176.1; -.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1 1
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9032 MW; ADED6F0E266AD98E CRC64;

Query Match 100.0%; Score 59; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTCG 9
Db 28 WSPCSVTCG 36

RESULT 8
Q9U0P8 PRELIMINARY; PRT; 80 AA.
AC Q9U0P8;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
CN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RC STRAIN=M4;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ269957; CAB64237.1; -.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1 1
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9073 MW; A8F404B8FB142B1E CRC64;

Query Match 100.0%; Score 59; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSPCSVTCG 9
Db 28 WSPCSVTCG 36

RESULT 9
Q9U0P7 PRELIMINARY; PRT; 80 AA.
AC Q9U0P7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
CN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
```

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-DA4230;
RA  de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT  "Sequence variation in the non-repeat region of the Plasmodium
RT  falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT  Burmese field isolates and from laboratory strains.";
RL  Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AJ269961; CAB64180.1; -.
DR  InterPro; IPR000884; TSP1.
DR  Pfam; PF00090; tsp_1; 1.
DR  PRINTS; PR01303; CRCMSPRZOITE.
DR  SMART; SM00209; TSP1; 1.
FT  NON_TER 1
FT  NON_TER 80
SQ  SEQUENCE 80 AA; 9047 MW; BA769C90DB031C3E CRC64;

Query Match 100.0%; Score 59; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
DB 28 WSPCSVTCG 36

RESULT 10
Q9U0P6 PRELIMINARY; PRT; 80 AA.
AC Q9U0P6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created).
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RC STRAIN-DA4264;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269963; CAB64182.1; -.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9002 MW; 1CEBAE08E6C9E976 CRC64;

Query Match 100.0%; Score 59; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
DB 28 WSPCSVTCG 36

RESULT 11
Q9U0P5 PRELIMINARY; PRT; 80 AA.
AC Q9U0P5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
```

```

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RC STRAIN-DA363;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269965; CAB64184.1; -.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9000 MW; 03798BD18F0BF3B3 CRC64;

Query Match 100.0%; Score 59; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
DB 28 WSPCSVTCG 36

RESULT 12
Q9U0P4 PRELIMINARY; PRT; 80 AA.
AC Q9U0P4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RC STRAIN-DA396;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269967; CAB64186.1; -.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 8988 MW; 0E7689D18F031B53 CRC64;

Query Match 100.0%; Score 59; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
DB 28 WSPCSVTCG 36
```

Query Match 100.0%; Score 59; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
DB 28 WSPCSVTCG 36

RESULT 15

OSU0P1 PRELIMINARY; PRT; 80 AA.
ID Q9U0P1
AC Q9U0P1; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B1896;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269978; CAB64197.1; -
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1;
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9102 MW; BFC6C970CEFOFA3E CRC64;

Query Match 100.0%; Score 59; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
DB 28 WSPCSVTCG 36

Search completed: January 29, 2002, 11:12:10
Job time: 766 sec

RESULT 13

OSU0P3 PRELIMINARY; PRT; 80 AA.
ID Q9U0P3
AC Q9U0P3; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D4405;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269969; CAB64188.1; -
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1;
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9033 MW; BA71EBE0DB03193B CRC64;

Query Match 100.0%; Score 59; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTCG 9
DB 28 WSPCSVTCG 36

RESULT 14

OSU0P2 PRELIMINARY; PRT; 80 AA.
ID Q9U0P2
AC Q9U0P2; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D4416;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269970; CAB64189.1; -
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1;
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9015 MW; 1CF404B8FB142C73 CRC64;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:21:43 ; Search time 310.82 Seconds
(without alignments)
2.145 Million cell updates/sec

Title: US-09-763-397A-9
Perfect score: 50
Sequence: 1 KPIVQYDNF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101.*

- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
- 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	9	AA43243	Cytotoxic T lympho
2	50	100.0	9	AA87311	Plasmodium falcipa
3	50	100.0	9	AA78844	P. falciparum live
4	50	100.0	9	AA03663	Amino acid sequenc
5	50	100.0	9	AA23658	Cytotoxic T lympho
6	50	100.0	9	AA70285	Plasmodium falcipa
7	50	100.0	229	AAV03681	Amino acid sequenc
8	50	100.0	350	AAV70278	Recombinant vaccin
9	50	100.0	493	AA26944	P. falciparum USA g
10	38	76.0	215	AA33138	Zea mays protein f
11	38	76.0	239	AA33137	Zea mays protein f

12	38	76.0	267	21	AA33136	Zea mays protein f
13	38	76.0	459	22	AA82076	S. epidermidis ope
14	38	76.0	459	22	AA82730	S. epidermidis ope
15	34	68.0	103	21	AA54615	Zea mays protein f
16	34	68.0	112	21	AA54417	Zea mays protein f
17	34	68.0	157	21	AA54612	Zea mays protein f
18	34	68.0	161	21	AA54415	Zea mays protein f
19	34	68.0	264	11	AA06464	Derived protein fr
20	34	68.0	264	20	AAV49850	Human pancreatic i
21	34	68.0	264	20	AAV33356	Human islet cell a
22	34	68.0	264	20	AAW80482	Islet cell antibod
23	34	68.0	284	21	AA49334	Pancreatic islet c
24	34	68.0	1490	22	AAW65644	Novel protein kina
25	33	66.0	24	22	AAW80922	Smad interacting p
26	33	66.0	64	20	AAV38898	Neisseria meningit
27	33	66.0	169	21	AAW07813	Amino acid sequenc
28	33	66.0	180	21	AAV28428	Wheat branched cha
29	33	66.0	342	21	AAV74492	Neisseria meningit
30	33	66.0	342	21	AAV74493	Neisseria meningit
31	33	66.0	349	20	AAV38900	N. meningitidis st
32	33	66.0	353	21	AAW07821	Amino acid sequenc
33	33	66.0	354	21	AAW07822	Amino acid sequenc
34	33	66.0	407	20	AAV38899	Neisseria meningit
35	33	66.0	1190	17	AAW95241	Human EHOC-1. Hom
36	32	64.0	175	22	AAW01053	CPE 56 protein seq
37	32	64.0	427	20	AAV08218	Staphylococcus cap
38	32	64.0	434	20	AAV08215	Staphylococcus hae
39	32	64.0	521	17	AAW14445	CarA gene product.
40	32	64.0	573	21	AAW16603	Bacteriophage 192
41	32	64.0	585	17	AAW96246	Malic enzyme #1.
42	32	64.0	1438	21	AAW18270	Plasmodium falcipa
43	31	62.0	215	21	AAW87855	Murine FGF-8 prote
44	31	62.0	216	22	AAU03676	Group B Streptococ
45	31	62.0	235	18	AAW15763	Cotton fibrous tis

ALIGNMENTS

RESULT 1	
AA43243	
ID	AA43243 standard; peptide; 9 AA.
XX	AA43243;
XX	04-MAY-1994 (first entry)
DT	
XX	
DE	Cytotoxic T lymphocyte recognition/induction peptide.
XX	
KW	CTL; vaccine; malaria; specific antigen-derived.
XX	
OS	Synthetic.
XX	
PN	W09320103-A.
XX	
PD	14-OCT-1993.
XX	
PF	05-APR-1993; 93WO-GB00711.
XX	
PR	03-APR-1992; 92GB-0008068.
PR	20-AUG-1992; 92GB-0017704.
XX	
PA	(ISIS-) ISIS INNOVATION LTD.
XX	
PI	Elvin J, Gotch FM, Hill AV, McMichael AJ, Whittle HC;
XX	
DR	WPI; 1993-336833/42.
XX	
PT	Peptide(s) recognising or inducing cytotoxic T lymphocytes -
PT	useful in vaccines against malaria or HIV-2, derived from
PT	specific antigen and human leukocyte antigen contg. class I
PT	restricted epitope
XX	

PS Claim 6; Page 30; 35pp; English.

XX The sequence is that of peptide Is6 which is recognised by, or can
CC induce, cytotoxic T lymphocytes. It may be useful in vaccines against
CC malaria.

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 50; DB 14; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9
DB 1 kplivqydnf 9

RESULT 2

AAAR87311
ID AAR87311 standard; peptide; 9 AA.

XX AC AAR87311;

XX DT 16-MAY-1996 (first entry)

XX DE Plasmodium falciparum HLA-B7 potential epitope Is6.

XX KW Liver-stage antigen; LSA-1; human leucocyte antigen; HLA; class 1;

XX KW HLA-B7; potential epitope; malaria; vaccine; CTL induction;

XX KW cytotoxic T lymphocyte.

XX OS Plasmodium falciparum.

XX PN WO9526982-A2.

XX PD 12-OCT-1995.

XX PF 31-MAR-1995; 95WO-GB00737.

XX PR 31-MAR-1994; 94GB-0006492.

XX PA (ISIS-) ISIS INNOVATION LTD.

XX PI Aidoo M, Allsopp CEM, Hill AVS, Lalvani A, Plebanski M;

XX PI Whittle HC;

XX DR WPI; 1995-358584/46.

XX PT Plasmodium falciparum peptide(s) - useful in vaccine compositions
for immunising against malaria

XX PS Claim 11; Page 21; 23pp; English.

XX CC Potential P.falciparum HLA epitopes are peptides derived from
antigens shown to be targets of CTL recognition and which are
capable of binding with high affinity to a specified HLA class 1
molecule. Although they have not been shown to be recognised by CTL
grown from malaria-exposed individuals their ability to bind to the
relevant class 1 molecule indicates that they are likely to be
presented on the surface of malaria-infected hepatocytes in vivo.
The present sequence is a potential HLA-B7 epitope derived from
liver-stage antigen-1. The use of this peptide (or its variants or
longer peptides comprising it) as a CTL inducer for immunisation
against malaria in HLA-B7 individuals is claimed.

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 50; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9
DB 1 kplivqydnf 9

RESULT 3

AAAR78844
ID AAR78844 standard; peptide; 9 AA.

XX AC AAR78844;

XX DT 27-MAR-1996 (first entry)

XX DE P. falciparum liver Ag 1786-1794 cytotoxic T lymphocyte epitope.

XX KW Liver Ag 1786-1794; cytotoxic T; CTL; epitope; helper T; HTL;

XX KW lymphocyte; viruses; parasites; tumours; antigens; treatment;

XX KW disease prevention; cell.

XX OS Plasmodium falciparum.

XX PN WO9522317-A1.

XX PD 24-AUG-1995.

XX PF 16-FEB-1995; 95WO-US02121.

XX PR 16-FEB-1994; 94US-0197484.

XX PA (CYTE-) CYTEL CORP.

XX PI Cent RW, Grey H, Sette AD, Vitiello MA;

XX DR WPI; 1995-302545/39.

XX PT Compsn. inducing cytotoxic T lymphocyte response to pref. viral,
bacterial, parasitic or tumour antigens - useful in the treatment
and prevention of diseases associated with the antigen e.g.

XX PT hepatitis B

XX PS Disclosure; Page 17; 109pp; English.

XX CC A compsn. which induces a cytotoxic T lymphocyte (CTL) response to
an antigen (Ag) in a mammal comprises, a CTL Ag response inducing
CC peptide (i.e. AAR78824-R78853) and a lipid conjugated helper T cell
inducing peptide. The compsn. induces a CTL response to bacterial,
CC viral or tumour Ags, and is therefore useful in the treatment and
prevention of diseases associated with the Ag.

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 50; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9
DB 1 kplivqydnf 9

RESULT 4

AAAY03663
ID AAY03663 standard; peptide; 9 AA.

XX AC AAY03663;

XX DT 07-JUN-1999 (first entry)

XX DE Amino acid sequence of the malaria (M) string CTL epitope Ls6.

XX CD8+ T-cell; immune response; antigen; priming composition; CTL; epitope;
KW cytotoxic T lymphocyte; boosting; poxvirus vector; PVV; pathogen; tumour;

KW malaria; parasite; P. falciparum; viral; bacterial; parasitic; cancer;
 KW melanoma; HIV; breast; colon; vaccination.
 OS Plasmodium falciparum.
 XX
 PN W09856919-A2.
 XX
 PD 17-DEC-1998.
 XX
 XX 09-JUN-1998; 98WO-GB01681.
 XX
 XX 09-JUN-1997; 97GB-0011957.
 PR
 XX (ISIS-) ISIS INNOVATION LTD.
 PA
 XX Blanchard T, Gilbert SC, Hanke T, Hill AVS, McMichael AJ;
 PI Plebanski M, Schneider J, Smith GL;
 PI
 XX WPI; 1999-070325/06.
 DR N-PSDB; AAX29203.
 XX
 XX Generating CD8-positive T cell response to target antigen using
 PT recombinant poxvirus - for treating or preventing malaria and HIV
 PT infection, also epitope strings from Plasmodium and HIV
 PT
 XX
 PS Claim 38; Page 18; 85pp; English.
 XX
 XX The invention relates to methods and reagents for generating a
 CC protective CD8+ T-cell immune response against at least one target
 CC antigen. The kits of the invention comprises (i) as priming composition,
 CC a source of one or more CD8+ T-cell [cytotoxic T lymphocytes-(CTL)]
 CC epitopes of the target antigen, plus a carrier and (ii) as boosting
 CC composition a source of CTL epitopes, with at least one CTL epitope the
 CC same as used in (i), with this source being a non-replicating or
 CC replication-impaired recombinant poxvirus vector (PVV) plus a carrier. If
 CC the source of CTL epitopes in (i) is a viral vector, then the vector in
 CC (ii) is from a different virus. The kits are used to generate an immune
 CC response (prophylactic or therapeutic) against pathogens or tumours,
 CC specifically against malaria parasites such as P. falciparum, or HIV, and
 CC also many other bacterial, viral or parasitic pathogens. The kits are
 CC also used for protective response against melanoma and cancer of breast
 CC or colon, and generally wherever a strong CD8+ response is protective.
 CC The boosting composition may be used alone to boost a naturally primed
 CC response against malaria. The specified PVV provide an excellent booster
 CC effect, better than that from wild-type poxvirus, resulting in complete
 CC rather than partial protection against sporozoite challenge. Also PVV are
 CC safer to use than wild-type virus. Sequences AAY03661-680 represent CTL
 CC peptide epitopes of the malaria (M) string.
 XX
 SQ Sequence 9 AA;

Query Match 100.0%; Score 50; DB 20; Length 9;
 Best Local Similarity 100.0%; Pred. No. 4.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KPIVQYDNF 9
 |||||
 Db 1 kplvqydnf 9

RESULT 5
 AAB23658
 ID AAB23658 standard; Peptide; 9 AA.
 XX
 AC AAB23658;
 XX
 XX 05-JAN-2001 (first entry)
 DT
 XX Cytotoxic T lymphocyte (CTL) epitope SEQ ID NO:10.
 DE
 XX
 KW ATPase; Hsp70; heat shock protein; cytotoxic T lymphocyte; CTL;
 KW immune response; infectious disease; malaria; cytotoxic T cell;

KW cytostatic; immunostimulant; cellular immune response inducer;
 KW protozoacide; leukaemia; cancer.
 OS Homo sapiens.
 XX
 PN W0200049041-A1.
 XX
 PD 24-AUG-2000.
 XX
 XX 18-FEB-2000; 2000WO-JP00941.
 PF
 XX 19-FEB-1999; 99JP-0041535.
 PR
 XX (SUME) SUMITOMO ELECTRIC IND CO.
 PA
 XX Shinbara N, Udonon H, Yui K;
 PI
 XX WPI; 2000-543748/49.
 DR
 XX Fused protein capable of inducing cellular immune response, useful as
 PT active ingredient for drug compositions in preventing and/or treating
 PT infectious diseases such as malaria or cancer
 PT
 XX Claim 5; Page 53; 72pp; Japanese.
 PS
 XX The present invention describes a fused protein (I) prepared from a
 CC peptide containing a CTL (cytotoxic T lymphocyte) epitope recognised by
 CC cytotoxic T cells and a protein containing the ATPase domain of a heat
 CC shock protein. Also described are: (1) a drug composition containing (1)
 CC as active ingredient; (2) a DNA encoding (1); (3) an expression vector
 CC containing the DNA of (2); and (4) a transformant which can retain the
 CC expression vector of (3). (1) has cytostatic, immunostimulant and
 CC protozoacide activities, and can be used as a cellular immune response
 CC inducer. The protein is useful as an active ingredient for drug
 CC compositions in preventing and/or treating infectious diseases such as
 CC malaria or cancer e.g. to provide systemic immunity against leukaemia.
 CC The present sequence represents a specifically claimed CTL epitope
 CC for use in a fused protein of the present invention.
 XX
 SQ Sequence 9 AA;

Query Match 100.0%; Score 50; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 4.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KPIVQYDNF 9
 |||||
 Db 1 kplvqydnf 9

RESULT 6
 AAY70285
 ID AAY70285 standard; peptide; 9 AA.
 XX
 AC AAY70285;
 XX
 XX 06-JUN-2000 (first entry)
 DT
 XX Plasmodium falciparum LSA-1 antigenic epitope, P595.
 DE
 XX Recombinant protein; CDC/NIAIDVAC-1; multivalent; malaria; vaccine;
 KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
 KW Circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
 KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-2;
 KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
 KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
 KW Pfg27; antiparasitic; prevention; anti-CDC/NIAIDVAC-1 antibody.
 XX
 OS Plasmodium falciparum.
 PN W0200011179-A1.
 XX


```

FT  /label= Mature_CDC/NIIMALVAC-1
FT  /note= *Recombinant multivalent malarial vaccine"
PN  WO200011179-A1.
PD  02-MAR-2000.
PP  19-AUG-1999; 99WO-US18869.
XX  21-AUG-1998; 98US-0097703.
PA  (NAIN-) NAT INST IMMUNOLOGY.
PA  (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX  Lal AA, Shi YP, Hasnain SE;
XX  WPI: 2000-237654/20.
XX  N-PSDB; AA251336.
XX  Novel recombinant protein as vaccine for treating malarial infection
PT  comprises antigenic peptides obtained from different stages of
PT  Plasmodium falciparum life cycle -
XX  Claim 3; Page 43-44; 52pp; English.
XX  The present sequence is that of recombinant protein CDC/NIIMALVAC-1,
CC  which is a multivalent, multistage malarial vaccine. The recombinant
CC  protein comprises, melittin signal peptide, (His)6 tag, T-cell epitope
CC  from tetanus toxoid and 21 antigenic epitopes from circumsporozoite
CC  protein (CSP), sporozoite surface protein-2 (SSP-2), liver stage
CC  antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1), MSP-2, apical
CC  membrane antigen-1 (AMA-1), erythrocyte binding antigen-175 (EBA-175),
CC  rhoptry associated protein-1 (RAP-1) and gamete specific antigen, Prg27.
CC  These epitopes were obtained at different stages of the life cycle of
CC  Plasmodium falciparum. CDC/NIIMALVAC-1 vaccine has antiparasitic
CC  activity and can be used for treatment and prevention of malarial
CC  infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for detecting
CC  P. falciparum in biological samples.
XX  Sequence 350 AA;

Query Match 100.0%; Score 50; DB 21; Length 350;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVQYDNF 9
DB 119 kpiqvdydnf 127
|||||||

RESULT 9
AAR26944
ID AAR26944 standard; Protein; 493 AA.
XX
AC AAR26944;
XX
DT 08-FEB-1993 (first entry)
XX
DE P.falciparum LSA gene C-terminal region.
XX
KW Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;
KW paludism; liver stage-specific antigen.
XX
OS Plasmodium falciparum.
XX
FH Key Location/Qualifiers
FT Region 13..213
FT Region /label= repetitive_region
FT Region 214..493
FT /label= non-repetitive_region
PN WO9213884-A.

```

```

XX 20-AUG-1992.
XX
XX 05-FEB-1992; 92WO-FR00104.
XX
XX 05-FEB-1991; 91FR-0001286.
XX
XX (INSP ) INST PASTEUR.
XX
XX Drulhe P, Guerin-Marchand C, Guerinmarchand C;
XX
XX WPI: 1992-299985/36.
XX N-PSDB; AAQ28119.
XX
XX Polypeptide(s) derived from liver stage of Plasmodium falciparum
PT - for vaccination against, treatment of and diagnosis of malaria
XX
XX Claim 2; Fig 8-10; 81pp; French.
XX
XX The 3' part of the P.falciparum liver-stage specific antigen (LSA)
CC gene codes for a polypeptide sequence which carries a T cell epitope
CC characteristic of a protein produced in hepatocytes infected with
CC P.falciparum. The polypeptide can be used in the preparation of
CC vaccines against malaria.
XX
XX Sequence 493 AA;

Query Match 100.0%; Score 50; DB 13; Length 493;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVQYDNF 9
DB 370 kpiqvdydnf 378
|||||||

RESULT 10
AAG33138
ID AAG33138 standard; Protein; 215 AA.
XX
AC AAG33138;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 40107.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
XX Zea mays subsp. mays.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 25-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.

```

us-09-763-397a-9.rag

Mon Feb 4 15:23:52 2002

PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139115.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140655.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144322.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.

PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0148171.
PR 10-AUG-1999; 99US-0148319.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148584.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157153.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159325.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161320.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.0%; Score 38; DB 21; Length 239;
Best Local Similarity 75.0%; Pred. No. 6.1;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVQYDN 8
||:||||
Db 106 kplvkyn 113

RESULT 12
AAG33136
ID AAG33136 standard; Protein; 267 AA.

XX AC AAG33136;

XX DT 18-OCT-2000 (first entry)

XX DE Zea mays protein fragment SEQ ID NO: 40105.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.

XX OS Zea mays subsp. mays.

XX PN EP1039405-A2.

XX PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

Mon Feb 4 15:23:52 2002

PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 76.0%; Score 38; DB 21; Length 267;
 Best Local Similarity 75.0%; Pred. No. 6.9;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDN 8
 ||:||||
 Db 134 kplvkyn 141

RESULT 13
 AAG82076
 ID AAG82076 standard; Protein; 459 AA.
 XX
 AC AAG82076;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE S. epidermidis open reading frame protein sequence SEQ ID NO:1246.
 XX
 DE Staphylococcus epidermidis SRI strain; infection; diagnosis;
 XX
 KW vaccination; endocarditis.
 KW
 XX Staphylococcus epidermidis.
 OS
 XX WO200134809-A2.
 PN
 XX 17-MAY-2001.
 PD
 XX 09-NOV-2000; 2000WO-US30782.
 PF
 XX 09-NOV-1999; 99US-0164258.
 PR
 XX (GLAX) GLAXO GROUP LTD.
 PA
 XX Kimmerly WJ;
 PI
 XX WPI; 2001-316495/33.
 DR
 XX N-PSDB; AAH52926.
 XX
 XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 PT useful for vaccinating against infections, e.g. endocarditis -
 PT
 XX Claim 18; Page 357; 2189pp; English.
 PS
 XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to
 CC AAH55098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention.
 CC N.B. the present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 XX
 XX Sequence 459 AA;

Query Match 76.0%; Score 38; DB 22; Length 459;
 Best Local Similarity 75.0%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PIVQYDNF 9
 ||:||||
 Db 343 plitydnf 350

RESULT 14
 AAG82730
 ID AAG82730 standard; Protein; 459 AA.
 XX
 AC AAG82730;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE S. epidermidis open reading frame protein sequence SEQ ID NO:2554.
 XX
 DE Staphylococcus epidermidis SRI strain; infection; diagnosis;
 XX
 KW vaccination; endocarditis.
 KW
 XX Staphylococcus epidermidis.
 OS
 XX WO200134809-A2.
 PN
 XX 17-MAY-2001.
 PD
 XX 09-NOV-2000; 2000WO-US30782.
 PF
 XX 09-NOV-1999; 99US-0164258.
 PR
 XX (GLAX) GLAXO GROUP LTD.
 PA
 XX Kimmerly WJ;
 PI
 XX WPI; 2001-316495/33.
 DR
 XX N-PSDB; AAH53580.
 XX
 XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 PT useful for vaccinating against infections, e.g. endocarditis -
 PT
 XX Claim 18; Page 670; 2188pp; English.
 PS
 XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to
 CC AAH55098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention.
 CC N.B. the present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 XX
 XX Sequence 459 AA;

Query Match 76.0%; Score 38; DB 22; Length 459;
 Best Local Similarity 75.0%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 PIVQYDNF 9
II: IIII
Db 343 plitydnf 350

RESULT 15
AAG54615
ID AAG54615 standard; Protein; 103 AA.
AC AAG54615;
XX
XX DT 18-OCT-2000 (first entry)
XX
XX DE Zea mays protein fragment SEQ ID NO: 69655.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.

Search completed: January 29, 2002, 10:21:44
Job time: 420 sec

PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 58.0%; Score 34; DB 21; Length 103;
Best Local Similarity 55.6%; Pred. No. 16;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPIVQYDNF 9
| | : | : |
Db 48 kpvlgyddf 56

THIS PAGE BLANK (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:24:03 ; Search time 133.18 Seconds
(without alignments)
1.521 Million cell updates/sec

Title: US-09-763-397A-9
Perfect score: 50
Sequence: 1 KPIVQYDNF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	9	2	US-08-318-856A-1
2	50	100.0	9	5	PCT-US95-02121-21
3	44	88.0	8	2	US-08-318-856A-46
4	34	68.0	264	2	US-08-468-576B-14
5	34	68.0	264	2	US-08-468-579B-14
6	34	68.0	264	3	US-08-468-577B-14
7	33	66.0	1190	1	US-08-337-650A-2
8	33	66.0	1190	4	US-08-048-887-2
9	32	64.0	521	2	US-08-737-825-3
10	31	62.0	235	2	US-08-580-545B-10
11	31	62.0	235	4	US-09-262-653A-10
12	31	62.0	282	4	US-09-412-102-2
13	31	62.0	282	4	US-09-217-787-2
14	31	62.0	285	4	US-09-412-102-4
15	31	62.0	285	4	US-09-217-787-4
16	30	60.0	30	3	US-08-433-522A-49
17	30	60.0	30	3	US-09-135-166-49
18	30	60.0	30	4	US-08-942-046-49
19	30	60.0	547	4	US-08-930-001-2
20	30	60.0	792	3	US-08-433-522A-8
21	30	60.0	792	3	US-08-135-166-8
22	30	60.0	792	4	US-08-942-046-8
23	30	60.0	793	3	US-08-433-522A-10
24	30	60.0	793	3	US-09-135-166-10
25	30	60.0	793	4	US-08-942-046-10
26	30	60.0	797	3	US-08-433-522A-2
27	30	60.0	797	3	US-08-433-522A-4

28	30	60.0	797	3	US-08-433-522A-6	Sequence 6, Appli
29	30	60.0	797	3	US-09-135-166-2	Sequence 2, Appli
30	30	60.0	797	3	US-09-135-166-4	Sequence 4, Appli
31	30	60.0	797	3	US-09-135-166-6	Sequence 6, Appli
32	30	60.0	797	4	US-08-942-046-2	Sequence 2, Appli
33	30	60.0	797	4	US-08-942-046-4	Sequence 4, Appli
34	30	60.0	797	4	US-08-942-046-6	Sequence 6, Appli
35	29	58.0	111	4	US-08-961-083-180	Sequence 180, App
36	29	58.0	142	2	US-08-448-438-1	Sequence 1, Appli
37	29	58.0	142	2	US-08-448-438-2	Sequence 2, Appli
38	29	58.0	142	2	US-08-448-438-3	Sequence 3, Appli
39	29	58.0	142	2	US-08-448-438-4	Sequence 4, Appli
40	29	58.0	157	3	US-08-992-176-5	Sequence 5, Appli
41	29	58.0	160	2	US-08-448-438-5	Sequence 5, Appli
42	29	58.0	160	2	US-08-448-438-6	Sequence 6, Appli
43	29	58.0	160	2	US-08-448-438-7	Sequence 7, Appli
44	29	58.0	160	2	US-08-448-438-8	Sequence 8, Appli
45	29	58.0	182	2	US-08-467-603-90	Sequence 90, Appli

ALIGNMENTS

RESULT 1
US-08-318-856A-1
; Sequence 1, Application US/08318856A
; Patent No. 5972351
; GENERAL INFORMATION:
; APPLICANT: Adrian V.S. Hill, et al;
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM MHC CLASS I-
; TITLE OF INVENTION: RESTRICTED CTL EPITOPES DERIVED FROM PRE-ERYTHROCYTIC STAGE
; TITLE OF INVENTION: ANTIGENS (AS AMENDED)
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,856A
; FILING DATE: October 3, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92 08 068.8
; FILING DATE: April 3, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92 17 704.7
; FILING DATE: August 20, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB93/00711
; FILING DATE: April 5, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: 263-PP1R15770US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 721-8200
; TELEFAX: (202) 721-8250
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum

US-08-318-856A-1

Query Match 100.0%; Score 50; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVQYDNF 9
|11111111|
DB 1 KPIVQYDNF 9

RESULT 2

PCT-US95-02121-21
; Sequence 21, Application PC/TUS9502121
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02121
; FILING DATE: 16-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,484
; FILING DATE: 16-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
PCT-US95-02121-21

Query Match 100.0%; Score 50; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVQYDNF 9
|11111111|
DB 1 KPIVQYDNF 9

RESULT 3

US-08-318-856A-46
; Sequence 46, Application US/08318856A

; Patent No. 5972351

; GENERAL INFORMATION:

; APPLICANT: Adrian V.S. Hill, et al.

; TITLE OF INVENTION: PLASMODIUM FALCIPARUM MHC CLASS I-

; TITLE OF INVENTION: RESTRICTED CTL EPITOPES DERIVED FROM PRE-ERYTHROCYTIC STAGE

; TITLE OF INVENTION: ANTIGENS (AS AMENDED)

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.

; STREET: 2033 K Street, N.W., Suite 800

; CITY: Washington

; STATE: D.C.S.A.

; COUNTRY: U.S.A.

; ZIP: 20006

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 mb

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WordPerfect 5.1+

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/318.856A

; FILING DATE: October 3, 1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 92 08 068.8

; FILING DATE: April 3, 1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 92 17 704.7

; FILING DATE: August 20, 1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/GB93/00711

; FILING DATE: April 5, 1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Lee Cheng

; REGISTRATION NUMBER: 40,949

; REFERENCE/DOCKET NUMBER: 263-PP1R1577US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 721-8200

; TELEFAX: (202) 721-8250

; INFORMATION FOR SEQ ID NO: 46:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 8 amino acid residues

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-318-856A-46

Query Match 88.0%; Score 44; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVQYDN 8
|1111111|
DB 1 KPIVQYDN 8

RESULT 4

US-08-468-576B-14
; Sequence 14, Application US/08468576B

; Patent No. 5955345

; GENERAL INFORMATION:

; APPLICANT: Rabin, Daniel

; TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS

; TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sprung Kramer Schaefer & Briscoe

; STREET: 660 White Plains Road

; CITY: Tarrytown

; STATE: New York

; COUNTRY: USA

; ZIP: 10591-5144

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: System 7.5
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,576B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,276
FILING DATE: 05-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,646
FILING DATE: 08-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/715,181
FILING DATE: 14-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/441,703
FILING DATE: 04-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/312,543
FILING DATE: 17-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/312,543
FILING DATE: 17-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: MDI 251.7-KGB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 264 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-468-576B-14

Query Match 68.0%; Score 34; DB 2; Length 264;
Best Local Similarity 55.6%; Pred. No. 24;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KPIVQYDNF 9
||: |||
Db 82 KPVTDYDQF 90

RESULT 5
US-08-468-579B-14
Sequence 14, Application US/08468579B
Patent No. 5981700
GENERAL INFORMATION:
APPLICANT: Rabin, Daniel
TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS
TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sprung Kramer Schaefer & Briscoe
STREET: 660 White Plains Road
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591-5144
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: System 7.5
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,579B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,276
FILING DATE: 05-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,646
FILING DATE: 08-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/715,181
FILING DATE: 14-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/441,703
FILING DATE: 04-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/312,543
FILING DATE: 17-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: MDI 251.5-KGB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 264 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-468-579B-14

Query Match 68.0%; Score 34; DB 2; Length 264;
Best Local Similarity 55.6%; Pred. No. 24;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KPIVQYDNF 9
||: |||
Db 82 KPVTDYDQF 90

RESULT 6
US-08-468-577B-14
Sequence 14, Application US/08468577B
Patent No. 6001804
GENERAL INFORMATION:
APPLICANT: Rabin, Daniel
TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS
TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sprung Kramer Schaefer & Briscoe
STREET: 660 White Plains Road
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591-5144
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: System 7.5
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,577B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,276
FILING DATE: 05-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,646
FILING DATE: 08-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/715,181
FILING DATE: 14-JUN-1991

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/441,703
;; FILING DATE: 04-DEC-1989
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/312,543
;; FILING DATE: 17-FEB-1989
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kurt G. Briscoe
;; REGISTRATION NUMBER: 33,141
;; REFERENCE/DOCKET NUMBER: MDI 251.8-KGB
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (914) 332-1700
;; TELEFAX: (914) 332-1844
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 264 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; US-08-468-577B-14

Query Match 68.0%; Score 34; DB 3; Length 264;
Best Local Similarity 55.6%; Pred. No. 24;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KPIVQYDNF 9
DB 82 KPIVTDYQF 90

RESULT 7
US-08-337-690A-2
; Sequence 2, Application US/08337690A
; Patent No. 5773268
; GENERAL INFORMATION:
; APPLICANT: Korenberg, Julie R.
; APPLICANT: Yamakawa, Kazuhiro
; TITLE OF INVENTION: A NOVEL CHROMOSOME 21 GENE MARKER.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS USING SAME
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL & FLORES, LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 91212
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 09-NOV-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-CE 2573
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)535-9001
; TELEFAX: (619)535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1190 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-337-690A-2

Query Match 66.0%; Score 33; DB 1; Length 1190;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PIVQYD 7
DB 774 PIVQYD 779

RESULT 8
US-09-048-987-2
; Sequence 2, Application US/09048887
; Patent No. 6166180
; GENERAL INFORMATION:
; APPLICANT: Korenberg, Julie R.
; APPLICANT: Yamakawa, Kazuhiro
; TITLE OF INVENTION: A NOVEL CHROMOSOME 21 GENE MARKER.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS USING SAME
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL & FLORES, LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 91212
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 09-NOV-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/337,690
; FILING DATE: 09-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-CE 2573
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)535-9001
; TELEFAX: (619)535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1190 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-048-887-2

Query Match 66.0%; Score 33; DB 4; Length 1190;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PIVQYD 7
DB 774 PIVQYD 779

RESULT 9
US-08-737-825-3
; Sequence 3, Application US/08737825
; Patent No. 5871922
; GENERAL INFORMATION:
; APPLICANT: SALMOND, GEORGE PEACOCK COPELAND
; APPLICANT: MCGOWAN, SIMON JAMES
; APPLICANT: SERAFIA, MOHAMMED
; APPLICANT: COX, ANTHONY RICHARD JOHN
; APPLICANT: HOLDEN, MATTHEW THOMAS GEOFFREY
; APPLICANT: PORTER, LAUREN ELIZABETH

APPLICANT: BYCROFT, BARRIE WALSHAM
APPLICANT: WILLIAMS, PAUL
APPLICANT: STEWART, GORDON SIDNEY ANDERSON BIRNIE
TITLE OF INVENTION: GENES INVOLVED IN THE BIOSYNTHETIC PATHWAY OF CARBAPENEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737,825
FILING DATE: 03-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 1009-0105P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 521 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Erwinia carotovora
US-08-737-825-3

Query Match 64.0%; Score 32; DB 2; Length 521;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KPIVOYDN 8
Db 374 KPGAQYDN 381

RESULT 10
US-08-580-545B-10
; Sequence 10, Application US/08580545B
; Patent No. 5932713
; GENERAL INFORMATION:
; APPLICANT: Yoshihisa, Kasukabe
; APPLICANT: Koichi, Fujisawa
; APPLICANT: Susumu, Nishiguchi
; APPLICANT: Yoshihiko, Maekawa
; APPLICANT: Randy, Allen
; TITLE OF INVENTION: COTTON FIBER TISSUE-SPECIFIC GENES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 601 Thirteenth Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: IBM PC compatible
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08580545B
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bretschneider, Barry E.
; REGISTRATION NUMBER: 28,055
; REFERENCE/DOCKET NUMBER: 04473/068001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/783-5070
; TELEFAX: 202/783-2331
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/580,545B
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bretschneider, Barry E.
REGISTRATION NUMBER: 28,055
REFERENCE/DOCKET NUMBER: 04473/068001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/783-5070
TELEFAX: 202/783-2331
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 235 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-580-545B-10

Query Match 62.0%; Score 31; DB 2; Length 235;
Best Local Similarity 71.4%; Pred. No. 83;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPIVOYD 7
Db 100 QPIVEYD 106

RESULT 11
US-09-262-653A-10
; Sequence 10, Application US/09262653A
; Patent No. 6166294
; GENERAL INFORMATION:
; APPLICANT: Yoshihisa, Kasukabe
; APPLICANT: Koichi, Fujisawa
; APPLICANT: Susumu, Nishiguchi
; APPLICANT: Yoshihiko, Maekawa
; APPLICANT: Randy, Allen
; TITLE OF INVENTION: COTTON FIBER TISSUE-SPECIFIC GENES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 601 Thirteenth Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/262,653A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bretschneider, Barry E.
; REGISTRATION NUMBER: 28,055
; REFERENCE/DOCKET NUMBER: 04473/068001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/783-5070
; TELEFAX: 202/783-2331
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-262-653A-10

Query Match 62.0%; Score 31; DB 4; Length 235;
Best Local Similarity 71.4%; Pred. No. 83;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYD 7
DB 100 QPIVEYD 106

RESULT 12
US-09-412-102-2
; Sequence 2, Application US/09412102
; Patent No. 6228992
; GENERAL INFORMATION:
; APPLICANT: JESSEN, HOLLY J
; APPLICANT: MEYER, TERRY E
; TITLE OF INVENTION: GENES AND METHODS FOR CONTROL OF
; NEMATODES IN PLANTS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave. Suite 310
; CITY: Raleigh
; STATE: NC
; COUNTRY: US
; ZIP: 27622
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/412,102
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/217,787
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-18P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 861 3175
; TELEFAX: 919 420 2202
; INFORMATION FOR SEQ ID NO: 2:
; LENGTH: 282 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-412-102-2

Query Match 62.0%; Score 31; DB 4; Length 282;
Best Local Similarity 57.1%; Pred. No. 1e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYD 7
DB 18 KPVLEYD 24

RESULT 13
US-09-217-787-2
; Sequence 2, Application US/09217787
; Patent No. 628498
; GENERAL INFORMATION:
; APPLICANT: JESSEN, HOLLY J
; APPLICANT: MEYER, TERRY E
; TITLE OF INVENTION: GENES AND METHODS FOR CONTROL OF

; TITLE OF INVENTION: NEMATODES IN PLANTS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave. Suite 310
; CITY: Raleigh
; STATE: NC
; COUNTRY: US
; ZIP: 27622
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/217,787
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-18P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 861 3175
; TELEFAX: 919 420 2202
; INFORMATION FOR SEQ ID NO: 2:
; LENGTH: 282 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-217-787-2

Query Match 62.0%; Score 31; DB 4; Length 282;
Best Local Similarity 57.1%; Pred. No. 1e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYD 7
DB 18 KPVLEYD 24

RESULT 14
US-09-412-102-4
; Sequence 4, Application US/09412102
; Patent No. 6228992
; GENERAL INFORMATION:
; APPLICANT: JESSEN, HOLLY J
; APPLICANT: MEYER, TERRY E
; TITLE OF INVENTION: GENES AND METHODS FOR CONTROL OF
; NEMATODES IN PLANTS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave. Suite 310
; CITY: Raleigh
; STATE: NC
; COUNTRY: US
; ZIP: 27622
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/412,102
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/217,787
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-18P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 420 2202
; TELEFAX: 919 881 3175
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 285 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-412-102-4

Query Match 62.0%; Score 31; DB 4; Length 285;
Best Local Similarity 57.1%; Pred. No. 1e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPIVOYD 7
||:|
Db 18 KPVLEYD 24

RESULT 15
US-09-217-787-4
; Sequence 4, Application US/09217787
; Patent No. 6284948
; GENERAL INFORMATION:
; APPLICANT: JESSEN, HOLLY J
; APPLICANT: MEYER, TERRY E
; TITLE OF INVENTION: GENES AND METHODS FOR CONTROL OF
; TITLE OF INVENTION: NEMATODES IN PLANTS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave. Suite 310
; CITY: Raleigh
; STATE: NC
; COUNTRY: US
; ZIP: 27622
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/217,787
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-18P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 420 2202
; TELEFAX: 919 881 3175
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 285 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-217-787-4

Query Match 62.0%; Score 31; DB 4; Length 285;
Best Local Similarity 57.1%; Pred. No. 1e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPIVOYD 7
||:|
Db 18 KPVLEYD 24

Search completed: January 29, 2002, 10:24:04
Job time: 510 sec

Mon Feb 4 15:23:52 2002

us-09-763-397a-9.ra1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:26:37 ; Search time 144.96 Seconds
(without alignments)
4.729 Million cell updates/sec

Title: US-09-763-397A-9
Perfect score: 50
Sequence: 1 KPVIQYDNF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68: *
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	1909	2 A45592	liver stage antigen
2	39	78.0	372	1 Q0BEJ7	UL100 protein - hu
3	38	76.0	558	2 S62458	vacuolar protein s
4	36	72.0	254	2 D82941	hypothetical prote
5	36	72.0	476	2 S09799	hypothetical prote
6	36	72.0	544	2 T27444	hypothetical prote
7	35	70.0	1071	2 B84062	hypothetical prote
8	34	68.0	129	2 D70153	hypothetical prote
9	34	68.0	160	2 D75388	NADH dehydrogenase
10	34	68.0	415	2 E71677	hypothetical prote
11	34	68.0	512	2 S74561	hypothetical prote
12	34	68.0	512	2 A70569	probable cpsA prot
13	34	68.0	1204	2 T19918	hypothetical prote
14	33.5	67.0	264	2 T10155	hypothetical prote
15	33	66.0	288	2 B72272	iron-sulfur cluste
16	33	66.0	350	2 T10498	UDPGlucose 4-epime
17	33	66.0	392	2 B82887	conserved hypothet
18	33	66.0	407	2 G81010	conserved hypothet
19	33	66.0	426	2 G38888	COI intron 14 prot
20	33	66.0	437	2 G82032	probable permease
21	33	66.0	519	2 T24784	hypothetical prote
22	33	66.0	688	2 T48176	receptor like prot
23	33	66.0	703	2 S25576	probable transport
24	33	66.0	703	2 S38400	mt2 protein - rat
25	33	66.0	725	2 S13426	multidrug resistan
26	33	66.0	747	2 T34329	hypothetical prote
27	33	66.0	906	2 E84948	NADH dehydrogenase
28	33	66.0	1259	2 Jc5523	transmembrane prot
29	33	66.0	1278	2 B22671	TyB protein - yeas

ALIGNMENTS

RESULT 1

A45592

liver stage antigen LSA-1 - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 22-Nov-1993 #sequence_revision 02-Dec-1994 #text_change 09-Jun-2000

C:Accession: S24597; A45592; S29393; S34842; B45592; C45592; D45592

R:Zhu, J.; Hollingdale, M. Library, November 1990

Submitted to the EMBL Data Bank

A:Reference number: S24597

A:Accession: S24597

A:Molecule type: DNA

A:Cross-references: EMBL:X56203; NID:g9915; PID:g9916

R:Zhu, J.; Hollingdale, M.R.

Mol. Biochem. Parasitol. 48, 223-226, 1991

A:Title: Structure of Plasmodium falciparum liver stage antigen-1.

A:Reference number: A45592; MUID:92107224

A:Accession: A45592

A:Molecule type: DNA

A:Residues: 1-195:638-688;1165-1215:1590-1909 <ZH2>

A:Note: sequence extracted from NCBI backbone (NCBIN:75010, NCBIN:75012, NCBIN:75014,

R:Guerin-Marchand, C.; Druilhe, P.; Galey, B.; Londono, A.; Patarapotikul, J.; Beaud

Nature 329, 164-167, 1987

A:Title: A liver-stage-specific antigen of Plasmodium falciparum characterized by gen

A:Reference number: S29393; MUID:87315391

A:Accession: S29393

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 323-387 <GUE1>

A:Cross-references: EMBL:M28266

R:Guerin-Marchand, C.; Druilhe, P.; Galey, B.; Londono, A.; Patarapotikul, J.; Beaud

Submitted to the EMBL Data Library, April 1992

A:Description: a liver-stage-sepcific antigen of plasmodium falciparum characterized

A:Reference number: S34842

A:Accession: S34842

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 323-381, 'HKAI' <GUE2>

A:Cross-references: EMBL:M28266

A:Note: difference at carboxyl end due to frameshift error

C:Comment: This protein is found as flocculent material in the parasitophorous vacuol

C:Superfamily: trichohyalin; calmodulin repeat homology

C:Keywords: EF hand

F:154-1629/Region: 17-residue repeats (A-K-E-K-L-Q-E-Q-S-D-L-E-Q-E-R-R)

Query Match 100.0%; Score 50; DB 2; Length 1909;

Best Local Similarity 100.0%; Pred. No. 0.09;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPVIQYDNF 9

|||||||

Db 1786 KPIVQYDNF 1794

RESULT 2

Q08E77
UL100 protein - human cytomegalovirus (strain AD169)
C:Species: human cytomegalovirus, human herpesvirus 5
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
C:Accession: S09865
R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.; M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A:Reference number: S09749; MUID:90269039
A:Accession: S09865
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-372 <CHE>
A:Cross-references: EMBL:X17403; NID:g59591; PIDN:CAA35336.1; PID:g1780879
A:Note: possible protein-coding frames are given
A:Note: the DNA sequence was submitted to EMBL, December 1989, in computer-readable form
C:Superfamily: cytomegalovirus UL100 protein

Query Match 78.0%; Score 39; DB 1; Length 372;
Best Local Similarity 87.5%; Pred. No. 2.7;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PIVQYDNF 9
Db 283 PIVQYDNF 290
|||||

RESULT 3

S62458
vacuolar protein sorting homolog-Sccl family - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 16-May-1996 #sequence_revision 13-Mar-1997 #text_change 10-Dec-1999
C:Accession: S62458; T38569
R:Badcock, K.; Churcher, C.M.
submitted to the EMBL Data Library, October 1995
A:Reference number: S62445
A:Accession: S62458
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-558 <BAD>
A:Cross-references: EMBL:Z54354; NID:g1019398; PIDN:CAA91168.1; PID:g1019401
R:Badcock, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, October 1995
A:Reference number: Z21745
A:Accession: T38569
A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA
A:Residues: 1-558 <BA2>
A:Cross-references: EMBL:Z54354; PIDN:CAA91168.1; GSPDB:GN00066; SPDB:SPAC2G11.03c
A:Experimental source: strain 972h; cosmid c2G11
C:Genetics:
A:Gene: SPDB:SPAC2G11.03c
A:Map position: 1L
A:Introns: 52/2; 121/3; 548/3
C:Superfamily: vacuolar protein sorting protein VPS45

Query Match 76.0%; Score 38; DB 2; Length 558;
Best Local Similarity 62.5%; Pred. No. 7;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPIVQYDN 8
Db 172 KPIVQYDN 179
|||||

RESULT 4

D82941
hypothetical protein U0031 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: D82941
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a
A:Reference number: A82870
A:Accession: D82941
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-254 <GLA>
A:Cross-references: GB:AS002103; GB:AF222894; NID:g6898977; PIDN:AAF30436.1; GSPDB:GN
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: U0031
A:Genetic code: SGC3

Query Match 72.0%; Score 36; DB 2; Length 254;
Best Local Similarity 66.7%; Pred. No. 7.5;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KPIVQYDNF 9
Db 51 KKLQVQYDNF 59
|||||

RESULT 5

S09799
hypothetical protein UL36 - human cytomegalovirus (strain AD169)
C:Species: human cytomegalovirus, human herpesvirus 5
A:Note: host Homo sapiens (man)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Oct-1999
C:Accession: S09799
R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.; M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A:Reference number: S09749; MUID:90269039
A:Accession: S09799
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-476 <CHE>
A:Cross-references: EMBL:X17403; NID:g59591; PIDN:CAA35335.1; PID:e298606; PID:g18139
A:Note: this sequence was submitted to the EMBL Data Library, December 1989
C:Genetics:
A:Introns: 68/1

Query Match 72.0%; Score 36; DB 2; Length 476;
Best Local Similarity 55.6%; Pred. No. 15;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPIVQYDNF 9
Db 142 RPMVQYDDY 150
|||||

RESULT 6

T27444
hypothetical protein Y7A9C.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27444
R:McMurray, A.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z20368
A:Accession: T27444
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA

A:Residues: 1-544 <WIL>
A:Cross-references: EMBL:Z99286; PIDN:CAB16542.1; GSPDB:GN00022; CESP:Y7A9C.1
A:Experimental source: clone Y7A9C
C:Genetics:
A:Gene: CESP:Y7A9C.1
A:Map position: 4
A:Introns: 27/1; 95/1; 296/3; 323/1; 391/3

Query Match 72.0%; Score 36; DB 2; Length 544;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KPIVQYDNF 9
| | | | |
Db 58 KPLVRIDNF 66

RESULT 7
B84062
Hypothetical protein BH3298 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
C:Accession: B84062
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20263314
A:Accession: B84062
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1071 <STO>
A:Cross-references: GB:AP001518; GB:BA000004; NID:gi0175792; PIDN:BA07017.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3298

Query Match 70.0%; Score 35; DB 2; Length 1071;
Best Local Similarity 75.0%; Pred. No. 62;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PIVQYDNF 9
| | | | |
Db 808 PELQYDNF 815

RESULT 8
D70153
Hypothetical protein BB0429 - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: D70153
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kralavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943
A:Accession: D70153
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-129 <KLE>
A:Cross-references: GB:AE001148; GB:AE000783; NID:g2680336; PIDN:AAC66810.1; PID:g268834
A:Experimental source: strain B31

Query Match 68.0%; Score 34; DB 2; Length 129;
Best Local Similarity 55.6%; Pred. No. 9;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KPIVQYDNF 9

Db 27 KPLEVDNF 35
| | | | |

RESULT 9
D75388
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 3 [similarity] - Deinococcus radio
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 19-May-2000
C:Accession: D75388
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: D75388
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-160 <WHI>
A:Cross-references: GB:AE001994; GB:AE000513; NID:g6459259; PIDN:AAF11072.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1506
A:Map position: 1
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 3
C:Keywords: membrane-associated complex; NAD; oxidoreductase

Query Match 68.0%; Score 34; DB 2; Length 160;
Best Local Similarity 77.8%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KPIVQYDNF 9
| | | | |
Db 39 KTIQYANF 47

RESULT 10
E71677
Hypothetical protein RP232 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C:Accession: E71677
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark
Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499
A:Accession: E71677
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-415 <AND>
A:Cross-references: GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAA14695.1; PID:g386
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: RP232
C:Superfamily: Rickettsia prowazekii hypothetical protein RP232

Query Match 68.0%; Score 34; DB 2; Length 415;
Best Local Similarity 85.7%; Pred. No. 34;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PIVQYDN 8
| | | | |
Db 206 PIVQHDN 212

RESULT 11
S74561
Hypothetical protein sll0225 - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S74561
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shmipo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* s.
A:Reference number: S74322; MUID:97061201
A:Accession: S74561
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-512 <KAN>
A:Cross-references: EMBL:D90900; GB:AB001339; NID:g1651766; PIDN:BAAI6713.1; PID:d101744
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Start codon: GTG

Query Match 68.0%; Score 34; DB 2; Length 512;
Best Local Similarity 62.5%; Pred. No. 43;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PIVQYDNF 9
1:1 |||
Db 83 PWSYDDF 90

RESULT 12

A70569
probable cpsA protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: A70569
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Hollroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987
A:Accession: A70569
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-512 <COL>
A:Cross-references: GB:295390; GB:AL123456; NID:g3261766; PIDN:CAB08707.1; PID:e315866;
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: cpsA

Query Match 68.0%; Score 34; DB 2; Length 512;
Best Local Similarity 71.4%; Pred. No. 43;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 PIVQYDN 8
1:1 |||
Db 331 PVRVYDN 337

RESULT 13

T19918
hypothetical protein C44B9.4 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19918
R:McMurray, A.
submitted to the EMBL Data Library, May 1996
A:Reference number: 219196
A:Accession: T19918
A:Status: Preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1204 <WIL>
A:Cross-references: EMBL:273424; PIDN:CAA97781.1; GSPDB:GN00021; CESP:C44B9.4

A:Experimental source: clone C44B9

C:Genetics:
A:Gene: CESP:C44B9.4
A:Map position: 3
A:Introns: 22/3; 66/3; 190/3; 226/2; 320/1; 426/1; 470/1; 687/1; 756/3; 806/1; 867/1;

Query Match 68.0%; Score 34; DB 2; Length 1204;
Best Local Similarity 75.0%; Pred. No. 11e02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PIVQYDNF 9
1:1 |||
Db 881 PAVYIDNF 888

RESULT 14

T10155
hypothetical protein D - phage T4
C:Species: phage T4
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 17-Nov-2000
C:Accession: T10155; T10156
R:Gram, H.; Ruger, W.
EMBO J. 4, 257-264, 1985
A:Title: Genes 55, alpha-gt, 47 and 46 of bacteriophage T4: the genomic organization
A:Reference number: A91016; MUID:85257446
A:Accession: T10155
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-264 <GRA>
A:Cross-references: EMBL:X01804; NID:g15229; PID:g15234
A:Accession: T10156
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 38-264 <GR2>
A:Cross-references: EMBL:X01804; NID:g15229; PID:g15235

Query Match 67.0%; Score 33.5; DB 2; Length 264;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 KPIVQYD---NF 9
||| ||| ||
Db 129 KPIVQYDLNGNF 140

RESULT 15

B72272
iron-sulfur cluster-binding protein - *Thermotoga maritima* (strain MSB8)
C:Species: *Thermotoga maritima*
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: B72272
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic Garrett, M.M.; Steward, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, C.M.

Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A:Reference number: A72200; MUID:99287316
A:Accession: B72272
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-288 <ARN>

A:Cross-references: GB:AE001784; GB:AE000512; NID:g4981842; PIDN:AAD36365.1; PID:g498
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1291
C:Superfamily: cell division inhibitor related protein; ferredoxin 2[4Fe-4S] homology

Query Match 66.0%; Score 33; DB 2; Length 288;
Best Local Similarity 62.5%; Pred. No. 36;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 KPIVOYDN 8
| | | | |
DB 262 KVVYTYN 269

Search Completed: January 29, 2002, 10:26:38
Job time: 649 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:13:39 ; Search time 80.65 Seconds
(without alignments)
4.092 Million cell updates/sec

Title: US-09-763-397A-9
Perfect score: 50
Sequence: 1 KP1VQYDNF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	39	78.0	372	1 VGLM_HCMVA	P16733 human cytom
2	38	76.0	558	1 YAB3_SCHPO	Q09805 schizosacch
3	36	72.0	476	1 UL36_HCMVA	P16767 human cytom
4	34	68.0	415	1 Y232_RICPR	Q92du2 rickettsia
5	34	68.0	1490	1 CRK7_HUMAN	Q9nyv4 homo sapien
6	33.5	67.0	264	1 Y03E_BPT4	P13329 bacterioph
7	33	66.0	261	1 YC33_MESVI	Q9mul6 mesostigma
8	33	66.0	350	1 GAE2_CYATE	Q65781 cyamopsis t
9	33	66.0	392	1 NADD_UREPA	Q9pq21 ureaplasma
10	33	66.0	519	1 YRV8_CAEEL	Q27520 caenorhabdi
11	33	66.0	703	1 TAP2_RAT	P36372 rattus norv
12	33	66.0	725	1 TAP1_RAT	P36370 rattus norv
13	33	66.0	906	1 NUOG_BUCAI	P57257 buchera ap
14	33	66.0	1156	1 JAK1_CYPCA	Q09178 cyprinus ca
15	33	66.0	1259	1 EH01_HUMAN	P48553 homo sapien
16	33	66.0	1755	1 Y277_YEAST	P47098 saccharomyc
17	32	64.0	235	1 XKDP_BACSU	P54335 bacillus su
18	32	64.0	249	1 HIS4_HAEIN	P44435 haemophilus
19	32	64.0	279	1 DEF6_DERFA	P49276 dermatophag
20	32	64.0	281	1 DEGV_BACSU	P32436 bacillus su
21	32	64.0	283	1 YD34_METJA	Q58730 methanococc
22	32	64.0	344	1 VGLM_HSV6U	Q04630 human herpe
23	32	64.0	344	1 VGLM_HSV6Z	P52449 human herpe
24	32	64.0	473	1 SCRB_LACLA	Q04937 lactococcus
25	32	64.0	586	1 YEJH_ECOLI	P33919 escherichia
26	32	64.0	659	1 VE1_HPV03	P36719 human papil
27	32	64.0	680	1 NCPR_CANNA	P50126 candida mal
28	32	64.0	1150	1 YC14_SCHPO	O74501 schizosacch
29	32	64.0	2179	1 POLG_EC23W	O73556 e genome po
30	32	64.0	2180	1 POLG_EC22H	Q66578 e genome po
31	31	62.0	118	1 RBS2_THIFE	Q07088 thioabacili
32	31	62.0	285	1 YW40_PSEAE	Q9hz00 pseudomonas
33	31	62.0	286	1 DAPA_CHLMU	Q9pk33 chlamydia m

RESULT 1				
VGLM_HCMVA	ID	VGLM_HCMVA	STANDARD;	PRT; 372 AA.
AC	P16733;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	GLYCOPROTEIN M.			
GN	GM OR UL100.			
OS	Human cytomegalovirus (strain AD169).			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Betaherpesvirinae; Cytomegalovirus.			
OX	NCBI_TaxID=10360;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89342616; PubMed=2547996;			
RA	Lehner R., Meyer H., Mach M.;			
RT	"Identification and characterization of a human cytomegalovirus gene			
RT	coding for a membrane protein that is conserved among human			
RT	herpesviruses.";			
RT	J. Virol. 63:3792-3800(1989).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90269039; PubMed=2161319;			
RA	Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,			
RA	Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,			
RA	Predie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;			
RT	"Analysis of the protein-coding content of the sequence of human			
RT	Cytomegalovirus strain AD169.";			
RL	Curr. Top. Microbiol. Immunol. 154:125-169(1990).			
CC	-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).			
CC	-!- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN M.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; M28350; AAA45984.1; -			
DR	EMBL; X17403; CAA35336.1; -			
DR	PIR; S09865; Q0BEJ7.			
DR	InterPro: IPR000785; Herpes_glycop.			
DR	Pfam; PF01528; Herpes_glycop; 1.			
DR	PRINTS; PR00333; HSVINTEGRMP.			
DR	Transmembrane; Glycoprotein.			
KW	Transmembrn 14 34			
FT	TRANSMEM 80 100			POTENTIAL.
FT	TRANSMEM 152 172			POTENTIAL.
FT	TRANSMEM 201 221			POTENTIAL.
FT	TRANSMEM 240 260			POTENTIAL.
FT	TRANSMEM 265 285			POTENTIAL.
FT	TRANSMEM 299 319			POTENTIAL.
FT	DOMAIN 359 372			ASP/GLU-RICH (ACIDIC).

084366 chlamydia t
P28321 saccharomyc
Q54430 streptococc
Q9pda6 xylella fas
O48456 klebsiella
P54376 bacillus su
P31921 euglena gra
Q27516 caenorhabdi
Q09653 caenorhabdi
Q92370 schizosacch
P38850 saccharomyc
P40824 salmonella

us-09-763-397a-9.rsp

Mon Feb 4 15:23:52 2002

```

FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 372 AA; 42861 MW;  F2F99BEC69BF8E32 CRC64;

Query Match 78.0%; Score 39; DB 1; Length 372;
Best Local Similarity 87.5%; Pred. No. 1.1;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PIVQYDNF 9
DB 283 PIVQYDTF 290

RESULT 2
ID YAB3_SCHPO STANDARD; PRT; 558 AA.
AC 009805;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PUTATIVE VACUOLAR PROTEIN SORTING-ASSOCIATED PROTEIN C2G11.03C.
GN SPAC2G11.03C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=972;
RC Badcock K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RA Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RL Submitter: BELONGS TO THE STXBP/UNC-18/SEC1 FAMILY. STRONG, TO
CC -1- SIMILARITY: BELONGS TO THE STXBP/UNC-18/SEC1 FAMILY. STRONG, TO
CC YEAST VPS45.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z54354; CAA91168.1; -
CC InterPro; IPR001619; Sec1.
CC Pfam; PF00995; Sec1; 1.
CC Hypothetical protein; Protein transport.
SQ SEQUENCE 558 AA; 63551 MW; 2C138B7CDF5CFEAF CRC64;

Query Match 76.0%; Score 38; DB 1; Length 558;
Best Local Similarity 62.5%; Pred. No. 2.8;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVQYDN 8
DB 172 KPIVRYDN 179

RESULT 3
ID UL36_HCMVA STANDARD; PRT; 476 AA.
AC P16767;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE HYPOTHETICAL PROTEIN UL36.
OS Human cytomegalovirus (strain AD169).
GN UL36.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;

```

```

RN SEQUENCE FROM N.A.
RP MEDLINE=90269039; PubMed=2161319;
RX Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,
RA Preddle E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
cytomegalovirus strain AD169."
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -1- SIMILARITY: BELONGS TO THE US22 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X17403; CAA35395.1; -
CC PIR; S09799; S09799.
CC InterPro; IPR003360; US22.
CC Pfam; PF02393; US22; 1.
CC Hypothetical protein.
SQ SEQUENCE 476 AA; 54981 MW; EDFFD4632CE0112C CRC64;

Query Match 72.0%; Score 36; DB 1; Length 476;
Best Local Similarity 55.6%; Pred. No. 6.3;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVQYDNF 9
DB 142 RPMVQYDDY 150

RESULT 4
ID Y232_RICPR STANDARD; PRT; 415 AA.
AC Q9ZD02;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN RP232.
GN RP232.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsia.
OX NCBI_TaxID=782;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O., Naeslund A.K.,
RA Sichert-Ponten T., Alsmark U.C.M., Podowski R.M., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
mitochondria."
RL Nature 396:133-140(1998).
CC -1- SIMILARITY: SOME, TO E.COLI YFGL.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ235271; CAA14695.1; -
CC InterPro; IPR002372; Bac_PQO-repeat.
CC Pfam; PF01011; Bacterial_PQO; 6.
CC Hypothetical protein; Complete proteome.
KW

```

SQ SEQUENCE 415 AA; 46750 MW; EF0A3604D5B302B3 CRC64;

Query Match 68.0%; Score 34; DB 1; Length 415;
Best Local Similarity 85.7%; Pred. No. 14;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 PIVQYDN 8
||||:|
Db 206 PIVQHDN 212

RESULT 5
ID CRK7_HUMAN STANDARD; PRT; 1490 AA.
AC Q9NV4; O94978;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CELL DIVISION CYCLE 2-RELATED PROTEIN KINASE 7 (EC 2.7.1.-) (CDC2-RELATED PROTEIN KINASE 7) (CKRKS).
GN CRK7 OR KIAA0904.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pines J.N., Kelly E.;
RT "CRK7: a novel CDC2-related protein kinase that colocalizes with interchromatin granule clusters";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 266-1262 FROM N.A.

RC TISSUE=Brain;
RX MEDLINE=99156230; PubMed=10048485;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.";
RL DNA Res. 5:355-364(1998).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF227198; AAF36401.1; -;
DR EMBL; AB020711; BAA74927.1; -;
DR HSSP; P27703; 2ERK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002965; P-rich_extensn.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Nuclear protein.
FT DOMAIN 727 1020 PROTEIN KINASE.
FT NP_BIND 733 741 ATP (BY SIMILARITY).
FT BINDING 756 756 ATP (BY SIMILARITY).
FT ACT_SITE 859 859 BY SIMILARITY.
FT DOMAIN 407 413 POLY-ALA.
FT DOMAIN 535 540 POLY-PRO.

FT DOMAIN 1266 1280 POLY-PRO.
FT CONFLICT 266 310 PGSTSRQSVSPPKPEKQAYQSSTSPSPYRRQRQSVSPYS
FT RRS -> SSSSRHSSISPVRLPLNSSLGAELSRKKKERA
FT AAAAAAKMGKE (IN REF. 2).
FT G -> D (IN REF. 2).
FT CONFLICT 639 639
FT CONFLICT 745 745 R -> K (IN REF. 2).
FT CONFLICT 1195 1195 M -> T (IN REF. 2).
FT CONFLICT 1254 1262 ACPPHLLP -> GRSNGNAL (IN REF. 2).
SQ SEQUENCE 1490 AA; 164154 MW; 851E18DF3BD2B1A1 CRC64;

Query Match 68.0%; Score 34; DB 1; Length 1490;
Best Local Similarity 62.5%; Pred. No. 60;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KPIVOYDN 8
|||:|
Db 68 KPLVEYDD 75

RESULT 6
Y03E_BPT4
ID Y03E_BPT4 STANDARD; PRT; 264 AA.
AC P13329;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE HYPOTHETICAL 30.4 KDA PROTEIN IN AGT-GP55 INTERGENIC REGION (ORF D).
GN Y03E OR AGT.1.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae;
OC T4-like phages.
OX NCBI_TaxID=10665;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85257446; PubMed=4018026;
RA Gram H., Rueger W.;
RT "Genes 55, alpha gt, 47 and 46 of bacteriophage T4: the genomic organization as deduced by sequence analysis";
RL EMBO J. 4:257-264(1985).
CC -1- SIMILARITY: TO PHAGE T4 MOBD AND Y04I.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X01804; CAA25938.1; -;
DR EMBL; X01804; CAA25939.1; ALT_INIT.
DR InterPro; IPR002711; HNH.
DR InterPro; IPR003615; HNH_nuc.
DR Pfam; PF01844; HNH; 1.
DR SMART; SM00507; HNHc; 1.
DR SMART; SM00497; IENR1; 2.
KW Hypothetical protein.
SQ SEQUENCE 264 AA; 30363 MW; 5121C882DC0DCE0 CRC64;

Query Match 67.0%; Score 33.5; DB 1; Length 264;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

Oy 1 KPIVOYD---NF 9
||| ||| ||
Db 129 KPIVOYDLNGNF 140

RESULT 7
YCX3_MESVI

```

ID YCX3 MESVI STANDARD; PRT; 261 AA.
AC O9MUL6;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE HYPOTHETICAL 30.4 KDA PROTEIN IN RPOZ-YCF20 INTERGENIC REGION.
OS Mesostigma viride.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Mesostigmatophyceae;
OC Mesostigmatiales; Mesostigmataceae; Mesostigma.
OX NCBI_TaxID=41882;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIES-296;
RX MEDLINE=20150907; PubMed=10668199;
RA Lemieux C., Otis C., Fournel M.;
RT "Ancestral chloroplast genome in Mesostigma viride reveals an early
RL branch of green plant evolution.";
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF156114; AAF3382.1;
KW Hypothetical protein; Chloroplast.
SQ SEQUENCE 261 AA; 30378 MW; F89E327F8810A074 CRC64;

Query Match 66.0%; Score 33; DB 1; Length 261;
Best Local Similarity 55.6%; Pred. No. 14;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 KPVOYDNP 9
DB 252 RPVOYSSY 260
:|||||:

RESULT 8
ID GAE2_CYATE STANDARD; PRT; 350 AA.
AC O65781;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE UDP-GLUCOSE 4-EPIMERASE GEPI48 (EC 5.1.3.2) (GALACTOWALDENASE) (UDP-
DE GALACTOSE 4-EPIMERASE).
OS Cyanopsis tetraonoloba (Guar) (Cluster bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Indigoferae;
OC Cyanopsis.
OX NCBI_TaxID=3832;
RN [1]
RP SEQUENCE FROM N.A.
RA Brunstedt J., Joerbo M., Pedersen S.G., Marcussen J.;
RT "Isolation and expression of two cDNA clones encoding UDP-galactose
RT epimerase genes expressed in guar endosperm.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: UDP-GLUCOSE = UDP-GALACTOSE.
CC -1- COFACTOR: NAD (BY SIMILARITY).
CC -1- PATHWAY: GALACTOSE METABOLISM AND BIOSYNTHESIS OF UDP-GALACTOSE AS
CC PRECURSOR OF GALACTOLIPIDS AND CELL WALL POLYSACCHARIDES.
CC -1- SIMILARITY: BELONGS TO THE SUGAR EPIMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ005082; CAA06339.1;
DR HSP; P09147; IXEL.
DR Mendel; 29884; Cyate; L265; 29884.
DR InterPro; IPR001509; Epimerase.
DR Pfam; PF01370; Epimerase; 1.
KW Isomerase; NAD; Galactose metabolism; Multigene family.
FT NP_BIND 5 36 NAD (POTENTIAL).
SQ SEQUENCE 350 AA; 38373 MW; C7E6C6CEBE7E31BC CRC64;

Query Match 66.0%; Score 33; DB 1; Length 350;
Best Local Similarity 62.5%; Pred. No. 19;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 KPVOYDN 8
DB 98 KPPLYDN 105
:|||||:

RESULT 9
ID NADL_UREPA STANDARD; PRT; 392 AA.
AC O9PQ21;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE PROBABLE NICOTINATE-NUCLEOTIDE ADENYLTRANSFERASE (EC 2.7.7.18)
DE (DEAMIDO-NAD(+)) PYROPHOSPHORYLASE (DEAMIDO-NAD(+)) DIPHOSPHORYLASE)
DE (NICOTINATE MONONUCLEOTIDE ADENYLTRANSFERASE) (NAMN
DE ADENYLTRANSFERASE).
GN NADD OR U0469.
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Ureaplasma.
OX NCBI_TaxID=134821;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROVAR 3; PubMed=11048724;
RX MEDLINE=20500219; Pubmed=11048724;
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA Cassell G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
RT urealyticum.";
RL Nature 407:757-762(2000).
CC -1- FUNCTION: CATALYZES THE REVERSIBLE ADENYLATION OF NICOTINATE
CC MONONUCLEOTIDE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + NICOTINATE RIBONUCLEOTIDE = DIPHOSPHATE
CC + DEAMIDO-NAD(+).
CC -1- PATHWAY: DE NOVO BIOSYNTHESIS OF NAD.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE NADD FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE IQEQ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE002143; AAF30881.1;
DR InterPro; IPR002819; HD.
DR InterPro; IPR003607; HDC.
DR Pfam; PF01966; HD; 1.
DR SMART; SM00471; HDC; 1.
KW Transferase; Nucleotidyltransferase; NAD; Complete proteome.
FT DOMAIN 1 210 NAMN ADENYLTRANSFERASE.
SQ SEQUENCE 392 AA; 46644 MW; 4897A34B2876725C CRC64;

```

Query Match 66.0%; Score 33; DB 1; Length 392;
 Best Local Similarity 55.6%; Pred. No. 21;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KPIVQYDNF 9
 Db 215 KHLISYDNF 223

RESULT 10
 YRV8_CAEEL STANDARD; PRT; 519 AA.
 AC Q27520;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PUTATIVE CYTOCHROME P450 CYP13A1 (EC 1.14.-.-).
 GN CYP13A1 OR T10B9.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Gardner A.;

RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
 CC MONOOXYGENASES. THEY OXIDIZE A VARIETY OF STRUCTURALLY UNRELATED
 CC COMPOUNDS, INCLUDING STEROIDS, FATTY ACIDS, AND XENOBIOTICS.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

EMBL: Z48717; CAA88610.1; -;
 DR WormPep; T10B9.8; CE01660.
 DR InterPro; IPR001128; Cyt_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PRINTS; PR00464; EP45011.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Hypothetical protein; Oxidoreductase; Monooxygenase; Heme.
 FT BINDING 465 465 HEME (BY SIMILARITY).
 SQ SEQUENCE 519 AA; 60044 MW; 3AF37EDC43539D7A CRC64;

Query Match 66.0%; Score 33; DB 1; Length 519;
 Best Local Similarity 71.4%; Pred. No. 30;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 IVQYDNF 9
 Db 92 VTQYDNF 98

RESULT 11
 TAP2_RAT STANDARD; PRT; 703 AA.
 AC P36372;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ANTIGEN PEPTIDE TRANSPORTER 2 (APT2).
 GN ABCB3 OR TAP2 OR MTP2.
 OS Rattus norvegicus (rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92100193; PubMed=1758495;
 RA Powis S.J.; Townsend A.R.M.; Deverson E.V.; Bastin J.;
 RA Butcher G.W.; Howard J.C.;
 RT "Restoration of antigen presentation to the mutant cell line RMA-S by
 RT an MHC-linked transporter.";
 RL Nature 354:528-531(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PVG-RTIL(LEW);
 RX MEDLINE=94266341; PubMed=8206525;
 RA Joly E.; Deverson E.V.; Coadwell W.L.; Guenther E.; Howard J.C.;
 RA Butcher G.W.;
 RT "The distribution of Tap2 alleles among laboratory rat RT1
 RT haplotypes.";
 RL Immunogenetics 40:45-53(1994).
 CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF ANTIGENS FROM THE CYTOPLASM
 CC TO A MEMBRANE-BOUND COMPARTMENT FOR ASSOCIATION WITH MHC CLASS I
 CC MOLECULES.
 CC -1- SUBUNIT: HETERODIMER OF TAP1 AND TAP2.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS). MDR SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

EMBL: X63854; CAA5339.1; -;
 DR EMBL; X75305; CAA53053.1; -;
 DR PIR; S19603; S19603.
 DR InterPro; IPR003593; AAA.
 DR InterPro; IPR003439; ABC_transportr.
 DR InterPro; IPR001140; ABC_transportr_tmem.
 DR InterPro; IPR001687; ATP_GTP_A.
 DR Pfam; PF00664; ABC_membrane; 1.
 DR Pfam; PF00005; ABC_tran; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 KW Peptide transport; Transport; Transmembrane; ATP-binding;
 KW Polymorphism.
 FT TRANSMEM 11 31 POTENTIAL.
 FT TRANSMEM 58 78 POTENTIAL.
 FT TRANSMEM 100 120 POTENTIAL.
 FT TRANSMEM 149 169 POTENTIAL.
 FT TRANSMEM 185 205 POTENTIAL.
 FT TRANSMEM 277 297 POTENTIAL.
 FT NP_BIND 503 510 ATP (POTENTIAL).
 FT NP_BIND 352 352 V -> F (IN TAP2L).
 FT VARIANT 603 603 K -> R (IN TAP2L).
 SQ SEQUENCE 703 AA; 77712 MW; E545993A8F784250 CRC64;

Query Match 66.0%; Score 33; DB 1; Length 703;
 Best Local Similarity 62.5%; Pred. No. 42;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPIVQYDN 8
 Db 532 EPLVQYDH 539

RESULT 12
 TAP1_RAT STANDARD; PRT; 725 AA.
 ID TAP1_RAT
 AC P36370;

Mon Feb 4 15:23:52 2002

DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ANTIGEN PEPTIDE TRANSPORTER 1 (APR1).
 GN ABCB2 OR TAP1 OR WTP1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91080926; PubMed=1979660;
 RA Deverson E.V.; Gow I.R., Coadwell W.J., Monaco J.J., Butcher G.W.,
 RT Howard J.C.;
 RL "MHC class II region encoding proteins related to the multidrug
 resistance family of transmembrane transporters";
 RL Nature 348:738-741(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HDX; TISSUE=Lymphocytes;
 RA Deverson E.V.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: INVOLVED IN THE TRANSPORT OF ANTIGENS FROM THE CYTOPLASM
 CC TO A MEMBRANE-BOUND COMPARTMENT FOR ASSOCIATION WITH MHC CLASS I
 CC MOLECULES.
 CC -!- SUBUNIT: HETERODIMER OF TAP1 AND TAP2.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS). MDR SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X57523; CAA40742.1; ALT_INIT.
 DR EMBL; Y10231; CAA71280.1; -.
 DR InterPro: IPR003593; AAA.
 DR InterPro: IPR003439; ABC_transporter.
 DR InterPro: IPR001140; ABC_transporter_tmam.
 DR InterPro: IPR001687; ATP_GTP_A.
 DR Pfam: PF00664; ABC_membrane; 1.
 DR Pfam: PF00005; ABC_tran; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 DR Peptide transport; Transport; Transmembrane; ATP-binding.
 KW TRANSMEM 68 88 POTENTIAL.
 FT TRANSMEM 117 137 POTENTIAL.
 FT TRANSMEM 164 184 POTENTIAL.
 FT TRANSMEM 205 225 POTENTIAL.
 FT TRANSMEM 298 318 POTENTIAL.
 FT TRANSMEM 392 412 POTENTIAL.
 FT TRANSMEM 421 441 POTENTIAL.
 FT NP_BIND 515 522 ATP (POTENTIAL).
 FT CARBOHYD 227 227 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 725 AA; 79150 MW; 3FA7215D0AC22EE0 CRC64;
 Query Match 66.0%; Score 33; DB 1; Length 725;
 Best Local Similarity 62.5%; Pred. No. 43;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPIVQYDN 8
 Db 544 EPLVQYDH 551
 RESULT 13
 NUOG_BUCAI STANDARD; PRT; 906 AA.
 ID NUOG_BUCAI
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ANTIGEN PEPTIDE TRANSPORTER 1 (APR1).
 GN ABCB2 OR TAP1 OR WTP1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91080926; PubMed=1979660;
 RA Deverson E.V.; Gow I.R., Coadwell W.J., Monaco J.J., Butcher G.W.,
 RT Howard J.C.;
 RL "MHC class II region encoding proteins related to the multidrug
 resistance family of transmembrane transporters";
 RL Nature 348:738-741(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HDX; TISSUE=Lymphocytes;
 RA Deverson E.V.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: INVOLVED IN THE TRANSPORT OF ANTIGENS FROM THE CYTOPLASM
 CC TO A MEMBRANE-BOUND COMPARTMENT FOR ASSOCIATION WITH MHC CLASS I
 CC MOLECULES.
 CC -!- SUBUNIT: HETERODIMER OF TAP1 AND TAP2.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS). MDR SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X57523; CAA40742.1; ALT_INIT.
 DR EMBL; Y10231; CAA71280.1; -.
 DR InterPro: IPR003593; AAA.
 DR InterPro: IPR003439; ABC_transporter.
 DR InterPro: IPR001140; ABC_transporter_tmam.
 DR InterPro: IPR001687; ATP_GTP_A.
 DR Pfam: PF00664; ABC_membrane; 1.
 DR Pfam: PF00005; ABC_tran; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 DR Peptide transport; Transport; Transmembrane; ATP-binding.
 KW TRANSMEM 68 88 POTENTIAL.
 FT TRANSMEM 117 137 POTENTIAL.
 FT TRANSMEM 164 184 POTENTIAL.
 FT TRANSMEM 205 225 POTENTIAL.
 FT TRANSMEM 298 318 POTENTIAL.
 FT TRANSMEM 392 412 POTENTIAL.
 FT TRANSMEM 421 441 POTENTIAL.
 FT NP_BIND 515 522 ATP (POTENTIAL).
 FT CARBOHYD 227 227 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 725 AA; 79150 MW; 3FA7215D0AC22EE0 CRC64;
 Query Match 66.0%; Score 33; DB 1; Length 725;
 Best Local Similarity 62.5%; Pred. No. 43;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPIVQYDN 8
 Db 544 EPLVQYDH 551
 RESULT 13
 NUOG_BUCAI STANDARD; PRT; 906 AA.
 ID NUOG_BUCAI
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ANTIGEN PEPTIDE TRANSPORTER 1 (APR1).
 GN ABCB2 OR TAP1 OR WTP1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91080926; PubMed=1979660;
 RA Deverson E.V.; Gow I.R., Coadwell W.J., Monaco J.J., Butcher G.W.,
 RT Howard J.C.;
 RL "MHC class II region encoding proteins related to the multidrug
 resistance family of transmembrane transporters";
 RL Nature 348:738-741(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HDX; TISSUE=Lymphocytes;
 RA Deverson E.V.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: INVOLVED IN THE TRANSPORT OF ANTIGENS FROM THE CYTOPLASM
 CC TO A MEMBRANE-BOUND COMPARTMENT FOR ASSOCIATION WITH MHC CLASS I
 CC MOLECULES.
 CC -!- SUBUNIT: HETERODIMER OF TAP1 AND TAP2.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS). MDR SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X57523; CAA40742.1; ALT_INIT.
 DR EMBL; Y10231; CAA71280.1; -.
 DR InterPro: IPR003593; AAA.
 DR InterPro: IPR003439; ABC_transporter.
 DR InterPro: IPR001140; ABC_transporter_tmam.
 DR InterPro: IPR001687; ATP_GTP_A.
 DR Pfam: PF00664; ABC_membrane; 1.
 DR Pfam: PF00005; ABC_tran; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 DR Peptide transport; Transport; Transmembrane; ATP-binding.
 KW TRANSMEM 68 88 POTENTIAL.
 FT TRANSMEM 117 137 POTENTIAL.
 FT TRANSMEM 164 184 POTENTIAL.
 FT TRANSMEM 205 225 POTENTIAL.
 FT TRANSMEM 298 318 POTENTIAL.
 FT TRANSMEM 392 412 POTENTIAL.
 FT TRANSMEM 421 441 POTENTIAL.
 FT NP_BIND 515 522 ATP (POTENTIAL).
 FT CARBOHYD 227 227 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 725 AA; 79150 MW; 3FA7215D0AC22EE0 CRC64;
 Query Match 66.0%; Score 33; DB 1; Length 906;
 Best Local Similarity 71.4%; Pred. No. 55;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 3 IVQYDNF 9
 Db 50 VTQYDNF 56
 RESULT 14
 JAK1_CYPCA STANDARD; PRT; 1156 AA.
 ID JAK1_CYPCA
 AC Q09178;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE TYROSINE-PROTEIN KINASE JAK1 (EC 2.7.1.112) (JANUS KINASE 1) (JAK-1).
GN JAK1
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprininae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97047914; PubMed=8892755;
RA Chang M.S., Chang G.D., Leu J.H., Huang F.L., Chou C.K., Huang C.J.,
RA Lo T.B.;
RT "Expression, characterization, and genomic structure of carp JAK1
RT kinase gene.";
RL DNA Cell Biol. 15:827-844(1996).
CC -!- FUNCTION: TYROSINE KINASE OF THE NON-RECEPTOR TYPE. INVOLVED IN
CC THE IFN-ALPHA/BETA/GAMMA SIGNAL PATHWAY. KINASE PARTNER FOR THE
CC INTERLEUKIN (IL)-2 RECEPTOR (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: WHOLLY INTRACELLULAR, POSSIBLY MEMBRANE
CC ASSOCIATED (BY SIMILARITY).
CC -!- DOMAIN: POSSESSES TWO PHOSPHOTRANSFERASE DOMAINS. THE SECOND ONE
CC PROBABLY CONTAINS THE CATALYTIC DOMAIN (BY SIMILARITY), WHILE THE
CC PRESENCE OF SLIGHT DIFFERENCES SUGGEST A DIFFERENT ROLE FOR
CC DOMAIN 1.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. JAK
CC SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: L24895; AAB38157.1; -.
DR EMBL: U53694; AAB38256.1; -.
DR EMBL: U53686; AAB38256.1; JOINED.
DR EMBL: U53687; AAB38256.1; JOINED.
DR EMBL: U53688; AAB38256.1; JOINED.
DR EMBL: U53689; AAB38256.1; JOINED.
DR EMBL: U53690; AAB38256.1; JOINED.
DR EMBL: U53691; AAB38256.1; JOINED.
DR EMBL: U53692; AAB38256.1; JOINED.
DR EMBL: U53693; AAB38256.1; JOINED.
DR HSP: P11362; IFC1.
DR InterPro: IPR000299; Band 4.1.
DR InterPro: IPR000719; Euk_Kinase.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001245; Tyr_kin.
DR Pfam: PF00069; Pkinase; 2.
DR SMART: SM00295; B41; 1.
DR SMART: SM00252; SH2; 1.
DR SMART: SM00219; TyrKc; 1.
DR PROSITE: P500107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: P500109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: P500111; PROTEIN_KINASE_DOM; 2.
DR PROSITE: P50001; SH2; FALSE_NEG.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW SH2 domain; Repeat
FT DOMAIN 337 359 LYS-RICH (BASIC).
FT DOMAIN 438 543 SH2 (ATYPICAL).
FT DOMAIN 593 857 PROTEIN KINASE 1.
FT DOMAIN 875 1153 PROTEIN KINASE 2.
FT NP_BIND 881 889 ATP (BY SIMILARITY).
FT BINDING 908 908 ATP (BY SIMILARITY).
FT ACT_SITE 1003 1003 BY SIMILARITY.
FT MOD_RES 1034 1034 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 1156 AA; 132245 MW; FF54BDD02F9666F5 CRC64;

Query Match 56.0%; Score 33; DB 1; Length 1156;
Best Local Similarity 55.6%; Pred. No. 73;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 KPIVQYDNF 9
:|: || ||
Db 489 RPVPQYKNE 497

RESULT 15
EHOI_HUMAN STANDARD; PRT; 1259 AA.
ID AC P48553; Q9Y4L3; Q9UND4;
DT 01-FEB-1996 (Rel. 33, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE EPILEPSY HOLOPROSENCEPHALY CANDIDATE-1 PROTEIN (EHOC-1) (TRANSMEMBRANE
DE PROTEIN 1) (GT334 PROTEIN).
GN TMEM1 OR EHOI1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=95359979; PubMed=7633421;
RA Yamakawa K., Mitchell S., Hubert R., Chen X.-N., Colbern S.,
RA Huo Y.-K., Gadomski C., Kim U.-J., Korenberg J.R.;
RT "Isolation and characterization of a candidate gene for progressive
RT myoclonous epilepsy on 21q22.3".
RL Hum. Mol. Genet. 4:709-716(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97339461; PubMed=9196060;
RA Nagamine K., Kudoh J., Kawasaki K., Minoshima S., Asakawa S., Ito F.,
RA Shimizu N.;
RT "Genomic organization and complete nucleotide sequence of the TMEM1
RT gene on human chromosome 21q22.3".
RL Biochem. Biophys. Res. Commun. 235:185-190(1997).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANTS GLU-257; MET-633 AND MET-726.
RX MEDLINE=98036064; PubMed=9370297;
RA Lafreniere R.G., Kibar Z., Rochefort D.L., Han F.-Y., Fon E.A.,
RA Dube M.-P., Kang X., Baird S., Korneluk R.G., Rommens J.M.,
RA Rouleau G.A.;
RT "Genomic structure of the human GT334 (EHOC-1) gene mapping to
RT 21q22.3".
RL Gene 198:313-321(1997).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED.
CC -!- DISEASE: CANDIDATE FOR AUTOIMMUNE POLYGLANDULAR DISEASE TYPE I
CC (APECED).
CC -!- SIMILARITY: BELONGS TO THE TMEM1 FAMILY.
CC -!- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 1187.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U19252; AAC50134.1; AUT_FRAME.
DR EMBL: AB001523; BAA21099.1; -.
DR EMBL: U61500; AAC51826.1; -.
DR EMBL: U61520; AAB58468.1; -.
DR EMBL: U61501; AAB58468.1; JOINED.
DR EMBL: U61502; AAB58468.1; JOINED.
DR EMBL: U61503; AAB58468.1; JOINED.
DR EMBL: U61504; AAB58468.1; JOINED.

DR EMBL; U61505; AAB58468.1; JOINED.
DR EMBL; U61506; AAB58468.1; JOINED.
DR EMBL; U61507; AAB58468.1; JOINED.
DR EMBL; U61508; AAB58468.1; JOINED.
DR EMBL; U61509; AAB58468.1; JOINED.
DR EMBL; U61510; AAB58468.1; JOINED.
DR EMBL; U61511; AAB58468.1; JOINED.
DR EMBL; U61512; AAB58468.1; JOINED.
DR EMBL; U61513; AAB58468.1; JOINED.
DR EMBL; U61514; AAB58468.1; JOINED.
DR EMBL; U61515; AAB58468.1; JOINED.
DR EMBL; U61516; AAB58468.1; JOINED.
DR EMBL; U61517; AAB58468.1; JOINED.
DR EMBL; U61518; AAB58468.1; JOINED.
DR EMBL; U61519; AAB58468.1; JOINED.
DR MIN: 602103; -
KW Polymorphism. 257 257 V -> E.
FT VARIANT 633 633 /FTID=VAR_009514.
FT VARIANT 633 633 I -> M.
FT VARIANT 726 726 /FTID=VAR_009515.
FT VARIANT 726 726 V -> M (FREQUENT POLYMORPHISM).
FT CONFLICT 114 114 /FTID=VAR_009516.
FT CONFLICT 121 121 V -> A (IN REF. 3; AAB58468).
FT CONFLICT 813 813 V -> A (IN REF. 3; AAB58468).
FT CONFLICT 910 910 Y -> D (IN REF. 1).
SQ SEQUENCE 1259 AA; 142188 MW; E8AB6847C9C6FE0C CRC64;
K -> Q (IN REF. 3).

Query Match 66.0%; Score 33; DB 1; Length 1259;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RIVQYD 7
Db 774 RIVQYD 779
|||||

Search completed: January 29, 2002, 11:13:40
Job time: 816 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:12:10 ; Search time 285.36 Seconds
(without alignments)
4.613 Million cell updates/sec

Title: US-09-763-397A-9

Perfect score: 50

Sequence: 1 KPVIQYDNF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL17.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	50	100.0	264	5 Q9GTX5	Q9GTX5 plasmodium
2	50	100.0	280	5 Q25843	Q25843 plasmodium
3	50	100.0	280	5 Q25844	Q25844 plasmodium
4	50	100.0	280	5 Q25845	Q25845 plasmodium
5	50	100.0	280	5 Q25846	Q25846 plasmodium
6	50	100.0	280	5 Q25847	Q25847 plasmodium
7	50	100.0	280	5 Q25848	Q25848 plasmodium
8	50	100.0	280	5 Q25849	Q25849 plasmodium
9	50	100.0	280	5 Q25850	Q25850 plasmodium
10	50	100.0	280	5 Q25851	Q25851 plasmodium
11	50	100.0	280	5 Q25852	Q25852 plasmodium
12	50	100.0	280	5 Q25853	Q25853 plasmodium
13	50	100.0	280	5 Q25854	Q25854 plasmodium
14	50	100.0	280	5 Q25855	Q25855 plasmodium
15	50	100.0	280	5 Q25889	Q25889 plasmodium
16	50	100.0	280	5 Q25900	Q25900 plasmodium
17	50	100.0	280	5 Q27243	Q27243 plasmodium
18	50	100.0	493	5 Q25886	Q25886 plasmodium
19	50	100.0	1909	5 Q25893	Q25893 plasmodium

20	47	94.0	280	5 Q25888	Q25888 plasmodium
21	36	72.0	254	2 Q9PRB4	Q9PRB4 ureaplasma
22	36	72.0	544	5 O18257	O18257 caenorhabdi
23	35	70.0	249	2 Q9RC01	Q9RC01 pseudomonas
24	35	70.0	1071	2 Q9K7R1	Q9K7R1 bacillus ha
25	34	68.0	94	9 Q9A292	Q9A292 lactobacill
26	34	68.0	129	2 O51390	O51390 borrelia bu
27	34	68.0	160	2 Q9RU86	Q9RU86 deinococcus
28	34	68.0	208	5 Q9VVM3	Q9VVM3 drosophila
29	34	68.0	373	11 Q9CVL4	Q9CVL4 mus musculu
30	34	68.0	512	2 P72706	P72706 synchocyst
31	34	68.0	512	2 O06347	O06347 mycobacteri
32	34	68.0	516	2 O50160	O50160 mycobacteri
33	34	68.0	1204	5 Q18605	Q18605 caenorhabdi
34	34	68.0	1266	5 Q9GRH4	Q9GRH4 spaeerechin
35	34	68.0	1399	5 Q9W427	Q9W427 drosophila
36	33	66.0	167	5 Q9NLMA	Q9NLMA leishmania
37	33	66.0	220	2 Q9F2H5	Q9F2H5 staphylococ
38	33	66.0	261	8 Q9MUL6	Q9MUL6 mesostigma
39	33	66.0	288	2 Q9X117	Q9X117 thermotoga
40	33	66.0	327	13 Q9PTQ9	Q9PTQ9 brachydanio
41	33	66.0	407	2 Q9JXF9	Q9JXF9 neisseria m
42	33	66.0	426	8 Q02691	Q02691 podospora a
43	33	66.0	437	2 Q9JW10	Q9JW10 neisseria m
44	33	66.0	496	10 Q9FR27	Q9FR27 lycopersico
45	33	66.0	688	10 Q9M020	Q9M020 arabidopsis

ALIGNMENTS

RESULT 1
Q9GTX5 PRELIMINARY: PRT; 264 AA.
AC Q9GTX5
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN-1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FCCL/HN;
RA Shan Z.X., Yu X.B., Li X.R., Ma C.L.;
RT "Molecular cloning and structure of the 3' terminal of liver stage
antigen-1 gene of Plasmodium falciparum isolate FCCL/HN";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF246996; FAG12324.1, -;
FT NON_TER
SQ SEQUENCE 264 AA: 31006 MW; 1714D653E8D9D1D7 CRC64;

Query Match 100.0%; Score 50; DB 5; Length 264;
Best Local Similarity 100.0%; Pred. No. 0.025; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

Qy 1 KPVIQYDNF 9
Db 141 KPVIQYDNF 149

RESULT 2
Q25843 PRELIMINARY: PRT; 280 AA.
AC Q25843
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.

OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
specific antigen-1 (LSA-1) of Plasmodium falciparum from field
isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40884; AAB59231.1; -
FT NON_TER 1
SQ SEQUENCE 280 AA; 32926 MW; 467080F32FAEAD33 CRC64;

Query Match 100.0%; Score 50; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPIVOYDNF 9
 |||||
Db 157 KPIVOYDNF 165

RESULT 3
Q25844
ID Q25844 PRELIMINARY; PRT; 280 AA.
AC Q25844;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
specific antigen-1 (LSA-1) of Plasmodium falciparum from field
isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40885; AAB59231.1; -
FT NON_TER 1
SQ SEQUENCE 280 AA; 32940 MW; E9708E3CFFAEAE9CF CRC64;

Query Match 100.0%; Score 50; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPIVOYDNF 9
 |||||
Db 157 KPIVOYDNF 165

RESULT 4
Q25845
ID Q25845 PRELIMINARY; PRT; 280 AA.
AC Q25845;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
specific antigen-1 (LSA-1) of Plasmodium falciparum from field
isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40886; AAB59232.1; -
FT NON_TER 1
SQ SEQUENCE 280 AA; 32908 MW; FDDE929F34DCA7E9 CRC64;

Query Match 100.0%; Score 50; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPIVOYDNF 9
 |||||
Db 157 KPIVOYDNF 165

RESULT 5
Q25846
ID Q25846 PRELIMINARY; PRT; 280 AA.
AC Q25846;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
specific antigen-1 (LSA-1) of Plasmodium falciparum from field
isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40887; AAC41597.1; -
FT NON_TER 1
SQ SEQUENCE 280 AA; 32882 MW; 46763641C81AFC33 CRC64;

Query Match 100.0%; Score 50; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPIVOYDNF 9
 |||||
Db 157 KPIVOYDNF 165

RESULT 6
Q25847
ID Q25847 PRELIMINARY; PRT; 280 AA.
AC Q25847;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

```
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40888; AAC41598.1; -
FT NON_TER 1
SQ SEQUENCE 280 AA; 32944 MW; 467085E23BABBC27 CRC64;

Query Match 100.0%; Score 50; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9
DB 157 KPIVOYDNF 165

RESULT 7
Q25848
ID Q25848 PRELIMINARY; PRT; 280 AA.
AC Q25848;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40889; AAC41599.1; -
FT NON_TER 1
SQ SEQUENCE 280 AA; 32939 MW; E77080385FAA0D61 CRC64;

Query Match 100.0%; Score 50; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9
DB 157 KPIVOYDNF 165

RESULT 8
Q25849
ID Q25849 PRELIMINARY; PRT; 280 AA.
AC Q25849;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40890; AAC41600.1; -
FT NON_TER 1
SQ SEQUENCE 280 AA; 32966 MW; E96D255154DEA9CF CRC64;
```

```
RP SEQUENCE FROM N.A.
RX STRAIN=KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40890; AAC41600.1; -
FT NON_TER 1
SQ SEQUENCE 280 AA; 32966 MW; E96D255154DEA9CF CRC64;

Query Match 100.0%; Score 50; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9
DB 157 KPIVOYDNF 165

RESULT 9
Q25850
ID Q25850 PRELIMINARY; PRT; 280 AA.
AC Q25850;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40891; AAC41601.1; -
FT NON_TER 1
SQ SEQUENCE 280 AA; 32989 MW; E96F812CFFAEBED8 CRC64;

Query Match 100.0%; Score 50; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9
DB 157 KPIVOYDNF 165

RESULT 10
Q25851
ID Q25851 PRELIMINARY; PRT; 280 AA.
AC Q25851;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40891; AAC41601.1; -
FT NON_TER 1
SQ SEQUENCE 280 AA; 32989 MW; E96F812CFFAEBED8 CRC64;
```

```
RC STRAIN-KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povo M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL: L40893; AAC41603.1; -.
FT NON_TER 1
SQ SEQUENCE 280 AA; 32927 MW; 48706EEDCF40ADD3 CRC64;

Query Match 100.0%; Score 50; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9
DB 157 KPIVOYDNF 165

RESULT 11
Q25852 PRELIMINARY; PRT; 280 AA.
AC Q25852;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povo M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL: L40834; AAB59233.1; -.
FT NON_TER 1
SQ SEQUENCE 280 AA; 32927 MW; E6648F85FBBAA245 CRC64;

Query Match 100.0%; Score 50; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9
DB 157 KPIVOYDNF 165

RESULT 12
Q25853 PRELIMINARY; PRT; 280 AA.
AC Q25853;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KENYAN;
```

```
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povo M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL: L40835; AAB59234.1; -.
FT NON_TER 1
SQ SEQUENCE 280 AA; 32938 MW; 46751C45F4DCBD33 CRC64;

Query Match 100.0%; Score 50; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9
DB 157 KPIVOYDNF 165

RESULT 13
Q25854 PRELIMINARY; PRT; 280 AA.
AC Q25854;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povo M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL: L40836; AAC41604.1; -.
FT NON_TER 1
SQ SEQUENCE 280 AA; 32966 MW; 61851C5AD6E1A211 CRC64;

Query Match 100.0%; Score 50; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9
DB 157 KPIVOYDNF 165

RESULT 14
Q25855 PRELIMINARY; PRT; 280 AA.
AC Q25855;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povo M.M.,
```

RA Hawley W.A., Collins W.E., Lal A.A.;
 RT "Sequence variations in the non-repetitive regions of the liver stage-
 RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
 RT isolates.";
 RL Mol. Biochem. Parasitol. 71:291-294(1995).
 DR EMBL: L40837; AAC41605.1; -
 FT NON_TER 1 1
 SQ SEQUENCE 280 AA; 32866 MW; 73164A709789DD31 CRC64;

Query Match 100.0%; Score 50; DB 5; Length 280;
 Best Local Similarity 100.0%; Pred. No. 0.027;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KPIVQYDNF 9
 |||||
 Db 157 KPIVQYDNF 165

RESULT 15

Q25889 PRELIMINARY; PRT: 280 AA.
 AC Q25889;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
 GN LSA-1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRAZILIAN;
 RX MEDLINE=96065765; PubMed=7477115;
 RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
 RA Hawley W.A., Collins W.E., Lal A.A.;
 RT "Sequence variations in the non-repetitive regions of the liver stage-
 RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
 RT isolates.";
 RL Mol. Biochem. Parasitol. 71:291-294(1995).
 DR EMBL: L40909; AAB59229.1; -
 FT NON_TER 1 1
 SQ SEQUENCE 280 AA; 32850 MW; FDEE929F34DF68E9 CRC64;

Query Match 100.0%; Score 50; DB 5; Length 280;
 Best Local Similarity 100.0%; Pred. No. 0.027;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KPIVQYDNF 9
 |||||
 Db 157 KPIVQYDNF 165

Search completed: January 29, 2002, 11:12:11
 Job time: 767 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:21:44 ; Search time 310.82 Seconds
(without alignments)
1.907 Million cell updates/sec

Title: US-09-763-397A-10
Perfect score: 44
Sequence: 1 KPNDKSLY 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101.*
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	8	14 AAR43244	Cytotoxic T lympho
2	44	100.0	8	16 AAR78842	P. falciparum live
3	44	100.0	8	19 AAW54665	Peptide from p.fal
4	44	100.0	8	20 AAY03661	Amino acid sequenc
5	44	100.0	8	21 AAB23657	Cytotoxic T lympho
6	44	100.0	8	21 AAB23657	Plasmodium falcipa
7	44	100.0	229	20 AAY03681	Amino acid sequenc
8	44	100.0	350	21 AAY70278	Recombinant vaccin
9	44	100.0	493	13 AAR26944	P.falciparum LSA g
10	36	81.8	514	22 AAU14727	Novel bone marrow
11	34	77.3	238	21 AAG28611	Arabidopsis thalia

12	34	77.3	260	21	AAG28610	Arabidopsis thalia
13	34	77.3	275	21	AAG28609	Arabidopsis thalia
14	34	77.3	380	19	AAW98800	H. pylori GHPO 128
15	34	77.3	903	22	AAW93106	Human protein sequ
16	33	75.0	663	13	AAW24101	Marek's disease vi
17	33	75.0	1522	20	AAW03183	Topoisomerase II b
18	32	72.7	272	22	AAW82359	S. epidermidis ope
19	32	72.7	330	22	AAW62569	Follicular conjunc
20	32	72.7	334	21	AAW54379	Arabidopsis thalia
21	32	72.7	341	21	AAW26206	Arabidopsis thalia
22	32	72.7	341	21	AAW41328	Arabidopsis thalia
23	32	72.7	341	21	AAW54378	Arabidopsis thalia
24	32	72.7	364	18	AAW22999	Canine herpesvirus
25	32	72.7	364	19	AAW72639	Canine herpesviru
26	32	72.7	364	22	AAW51316	Canine herpes viru
27	32	72.7	412	21	AAW26205	Arabidopsis thalia
28	32	72.7	412	21	AAW41327	Arabidopsis thalia
29	32	72.7	412	21	AAW54377	Arabidopsis thalia
30	32	72.7	423	21	AAW26204	Arabidopsis thalia
31	32	72.7	423	21	AAW41326	Arabidopsis thalia
32	32	72.7	435	22	AAW83046	S. epidermidis ope
33	32	72.7	967	20	AAW93621	Human CD13/aminope
34	32	72.7	977	21	AAW54345	Human pancreatic c
35	32	72.7	977	22	AAW24098	Human EST encoded
36	31	70.5	17	21	AAW26031	Human Ige C-epsilo
37	31	70.5	45	14	AAW31370	TAP 29 protein. T
38	31	70.5	431	22	AAW91713	C glutamic prote
39	31	70.5	443	19	AAW98460	H. pylori GHPO 716
40	31	70.5	553	21	AAW74579	Neisseria gonorrhe
41	31	70.5	553	21	AAW74580	Neisseria meningit
42	31	70.5	553	21	AAW74581	Neisseria meningit
43	30	68.2	59	22	AAW80361	Human prostate can
44	30	68.2	60	22	AAW80314	Human prostate can
45	30	68.2	63	21	AAW56455	Human prostate can

ALIGNMENTS

RESULT 1
AAR43244
ID AAR43244 standard; peptide; 8 AA.
XX
AC AAR43244;
XX
DT 04-MAY-1994 (first entry)
XX
DE Cytotoxic T lymphocyte recognition/induction peptide.
XX
KW CTL; vaccine; malaria; specific antigen-derived.
XX
OS Synthetic.
XX
PN WO9320103-A.
XX
PD 14-OCT-1993.
XX
PF 05-APR-1993; 93WO-GB00711.
XX
PR 03-APR-1992; 92GB-0008068.
PR 20-AUG-1992; 92GB-0017704.
XX
XX (ISIS-) ISIS INNOVATION LTD.
XX
XX Elvin J, Gotch FM, Hill AV, McMichael AJ, Whittle HC;
XX
XX WPI; 1993-336833/42.
XX
XX Peptide(s) recognising or inducing cytotoxic T lymphocytes -
XX
XX useful in vaccines against malaria or HIV-2, derived from
XX
XX specific antigen and human leukocyte antigen conty. class I
XX
XX restricted epitope

PS Claim 6; Page 30; 35pp; English.

CC The sequence is that of peptide ls8 which is recognised by, or can
CC induce, cytotoxic T lymphocytes. It may be useful in vaccines against
CC malaria.

XX Sequence 8 AA;

Query Match 100.0%; Score 44; DB 14; Length 8;

Best Local Similarity 100.0%; Pred. No. 4.3e+05;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KPNDKSLY 8

DB 1 kpndksly 8

RESULT 2

AAR78842 AAR78842 standard; peptide; 8 AA.

XX AC AAR78842;

XX DT 27-MAR-1996 (first entry)

XX DE P. falciparum liver Ag 1850-1857 cytotoxic T lymphocyte epitope.

XX KW Liver Ag 1850-1857; cytotoxic T; CTL; epitope; helper T; HTL;

XX KW lymphocyte; viruses; parasites; tumours; antigens; treatment;

XX KW disease prevention; cell.

XX OS Plasmodium falciparum.

XX PN WO9522317-A1.

XX PD 24-AUG-1995.

XX PF 16-FEB-1995; 95WO-US02121.

XX PR 16-FEB-1994; 94US-0197484.

XX PS (CYTE-) CYTEL CORP.

XX PI Ceut RW, Grey H, Sette AD, Vitello MA;

XX DR WPI; 1995-302545/39.

XX PT Compn. inducing cytotoxic T lymphocyte response to pref. viral,
XX bacterial, parasitic or tumour antigens - useful in the treatment
XX and prevention of diseases associated with the antigen e.g.
XX hepatitis B

PS Disclosure; Page 17; 109pp; English.

XX A compn. which induces a cytotoxic T lymphocyte (CTL) response to
XX an antigen (Ag) in a mammal comprises, a CTL Ag response inducing
XX peptide (i.e. AAR78842-R78853) and a lipid conjugated helper T cell
XX inducing peptide. The compn. induces a CTL response to bacterial,
XX viral or tumour Ags, and is therefore useful in the treatment and
XX prevention of diseases associated with the Ag.

XX Sequence 8 AA;

Query Match 100.0%; Score 44; DB 16; Length 8;

Best Local Similarity 100.0%; Pred. No. 4.3e+05;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KPNDKSLY 8

DB 1 kpndksly 8

RESULT 3

AAW54665

ID AAW54665 standard; peptide; 8 AA.

XX AC AAW54665;

XX DT 25-SEP-1998 (first entry)

XX DE Peptide from P. falciparum LS 1850-1857.

XX KW Mannose; antigen; antigen-presenting cell; mannoseylated peptide; T cell;

XX KW vaccine; treatment.

XX OS Synthetic.

XX PN WO9813378-A1.

XX PD 02-APR-1998.

XX PF 25-SEP-1997; 97WO-NL00536.

XX PR 26-SEP-1996; 96EP-0202701.

XX PA (UYLE-) RIJKSUNIV LEIDEN.

XX PI Drijfhout JW, Koning F;

XX DR WPI; 1998-230631/20.

XX PT Increasing uptake and presentation of antigen(s) - by adding mannose

XX PT residue(s) to antigen for increasing T cell response, useful in,

XX PT e.g. vaccines against viral infection(s)

XX PS Disclosure; Page 29; 47pp; English.

XX CC The peptides AAW5459-W54809 are examples of peptides to which at least
XX 1 (preferably 2) mannose can be attached to increase their uptake as
XX antigens by antigen-presenting cells. Uptake of agonist mannoseylated
XX peptides will increase the T cell response whereas uptake of antagonist
XX peptides blocks the T cell response. Blocking binding of immunogenic
XX autoantigens can be used in treatment of type I diabetes, rheumatoid
XX arthritis, graft rejection etc., also to induce T-cell non-
XX responsiveness. Vaccines containing mannoseylated antigen are used to
XX prevent or treat infections by, e.g. bacteria, viruses, fungi, helminths
XX and parasites.

XX SQ Sequence 8 AA;

Query Match 100.0%; Score 44; DB 19; Length 8;

Best Local Similarity 100.0%; Pred. No. 4.3e+05;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KPNDKSLY 8

DB 1 kpndksly 8

RESULT 4

AAAY03661

ID AAAY03661 standard; peptide; 8 AA.

XX AC AAAY03661;

XX DT 07-JUN-1999 (first entry)

XX DE Amino acid sequence of the malaria (M) string CTL epitope ls8.

XX CD8+ T-cell; immune response; antigen; priming composition; CTL; epitope;
XX cytotoxic T lymphocyte; boosting; poxvirus vector; PVV; pathogen; tumour;
XX malaria; parasite; P. falciparum; viral; bacterial; parasitic; cancer;
XX melanoma; HIV; breast; colon; vaccination.

```

XX OS Plasmodium falciparum.
XX PN WO9856919-A2.
XX PD 17-DEC-1998.
XX PF 09-JUN-1998; 98WO-GB01681.
XX PR 09-JUN-1997; 97GB-0011957.
XX PA (ISIS-) ISIS INNOVATION LTD.
XX PI Blanchard T, Gilbert SC, Hanke T, Hill AVS, McMichael AJ;
XX PI Plebanski M, Schneider J, Smith GL;
XX DR WPI: 1999-070325/06.
XX DR N-PSDB: AAX29201.
XX PT Generating CD8-positive T cell response to target antigen using
XX PT recombinant poxvirus - for treating or preventing malaria and HIV
XX PT infection, also epitope strings from Plasmodium and HIV
XX PS Claim 38; Page 18; 85pp; English.
XX CC The invention relates to methods and reagents for generating a
XX CC protective CD8+ T-cell immune response against at least one target
XX CC antigen. The kits of the invention comprises (i) as priming composition,
XX CC a source of one or more CD8+ T-cell [cytotoxic T lymphocytes-(CTL)]
XX CC epitopes of the target antigen, plus a carrier and (ii) as boosting
XX CC composition a source of CTL epitopes, with at least one CTL epitope the
XX CC same as used in (i), with this source being a non-replicating or
XX CC replication-impaired recombinant poxvirus vector (PVV) plus a carrier. If
XX CC the source of CTL epitopes in (i) is a viral vector, then the vector in
XX CC (ii) is from a different virus. The kits are used to generate an immune
XX CC response (prophylactic or therapeutic) against pathogens or tumours,
XX CC specifically against malaria parasites such as P. falciparum, or HIV, and
XX CC also many other bacterial, viral or parasitic pathogens. The kits are
XX CC also used for protective response against melanoma and cancer of breast
XX CC or colon, and generally wherever a strong CD8+ response is protective.
XX CC The boosting composition may be used alone to boost a naturally primed
XX CC response against malaria. The specified PVV provide an excellent booster
XX CC effect, better than that from wild-type poxvirus, resulting in complete
XX CC rather than partial protection against sporozoite challenge. Also PVV are
XX CC safer to use than wild-type virus. Sequences AAY03661-680 represent CTL
XX CC peptide epitopes of the malaria (M) string.
XX SQ Sequence 8 AA;

Query Match 100.0%; Score 44; DB 20; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
Db 1 kpndksly 8

RESULT 5
AAB23657
ID AAB23657 standard; Peptide; 8 AA.
XX AC AAB23657;
XX XX
XX DT 05-JAN-2001 (first entry)
XX DE Cytotoxic T lymphocyte (CTL) epitope SEQ ID NO:9.
XX KW ATPase; Hsp70; heat shock protein; cytotoxic T lymphocyte; CTL;
XX KW immune response; infectious disease; malaria; cytotoxic T cell;
XX KW cytotaxic; immunostimulant; cellular immune response inducer;
XX KW protozoacide; leukaemia; cancer.

```

```

XX OS Homo sapiens.
XX PN WO200049041-A1.
XX PD 24-AUG-2000.
XX PF 18-FEB-2000; 2000WO-JP00941.
XX PR 19-FEB-1999; 99JP-0041535.
XX PA (SUME) SUMITOMO ELECTRIC IND CO.
XX PI Shinbara N, Udono H, Yui K;
XX DR WPI: 2000-543748/49.
XX PT Fused protein capable of inducing cellular immune response, useful as
XX PT active ingredient for drug compositions in preventing and/or treating
XX PT infectious diseases such as malaria or cancer -
XX PS Claim 5; Page 53; 72pp; Japanese.
XX CC The present invention describes a fused protein (I) prepared from a
XX CC peptide containing a CTL (cytotoxic T lymphocyte) epitope recognised by
XX CC cytotoxic T cells and a protein containing the ATPase domain of a heat
XX CC shock protein. Also described are: (1) a drug composition containing (I)
XX CC as active ingredient; (2) a DNA encoding (I); (3) an expression vector
XX CC containing the DNA of (2); and (4) a transformant which can retain the
XX CC expression vector of (3). (I) has cytostatic, immunostimulant and
XX CC protozoacide activities, and can be used as a cellular immune response
XX CC inducer. The protein is useful as an active ingredient for drug
XX CC compositions in preventing and/or treating infectious diseases such as
XX CC malaria or cancer e.g. to provide systemic immunity against leukaemia.
XX CC The present sequence represents a specifically claimed CTL epitope
XX CC for use in a fused protein of the present invention.
XX SQ Sequence 8 AA;

Query Match 100.0%; Score 44; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
Db 1 kpndksly 8

RESULT 6
AAY70286
ID AAY70286 standard; peptide; 8 AA.
XX AC AAY70286;
XX XX
XX DT 06-JUN-2000 (first entry)
XX DE Plasmodium falciparum LSA-1 antigenic epitope, p596.
XX KW Recombinant protein; CDC/N1MALVAC-1; multivalent; malaria; vaccine;
XX KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
XX KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
XX KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
XX KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
XX KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
XX KW Pf927; antiparasitic; prevention; anti-CDC/N1MALVAC-1 antibody.
XX OS Plasmodium falciparum.
XX PN WO200011179-A1.
XX PD 02-MAR-2000.

```

PF 19-AUG-1999; 99NO-US18869.
 XX
 XX 21-AUG-1998; 98US-0097703.
 XX
 XX (NAIM-) NAT INST IMMUNOLOGY.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Lal AA, Shi YP, Hasnain SE;
 PI WPI: 2000-237654/20.
 XX
 XX Novel recombinant protein as vaccine for treating malarial infection
 PT comprises antigenic peptides obtained from different stages of
 PT plasmodium falciparum life cycle
 XX
 XX Claim 2; Page 17; 52pp; English.
 PS
 XX The present sequence is the antigenic epitope P596, derived from
 CC liver stage antigen-1 (LSA-1) of Plasmodium falciparum. It is used
 CC in the construction of recombinant protein CDC/NIIMALVAC-1, which
 CC is a multivalent, multistage malarial vaccine. The recombinant
 CC protein comprises, melittin signal peptide, (His)6 tag,
 CC T-cell epitope from tetanus toxoid and 21 antigenic epitopes from
 CC circumsporozoite protein (CSP), sporozoite surface protein-2 (SSP-2),
 CC liver stage antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1),
 CC MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding antigen-175
 CC (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete stages of the
 CC antigen, Pf27. These epitopes were obtained at different stages of the
 CC life cycle of P. falciparum. CDC/NIIMALVAC-1 vaccine has antiparasitic
 CC activity and can be used for treatment and prevention of malarial
 CC infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for detecting
 CC P. falciparum in biological samples.
 XX
 XX Sequence 8 AA;

Query Match 100.0%; Score 44; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 4.3e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KPNDSLSY 8
 |||||
 Db 1 Kpnksly 8

RESULT 7
 AAY03681
 ID AAY03681 standard; protein; 229 AA.

XX AC AAY03681;
 XX
 XX 07-JUN-1999 (first entry)
 XX
 XX Amino acid sequence of the malaria (M) string CTL epitope.

XX CD8+ T-cell; immune response; antigen; priming composition; CTL; epitope;
 KW cytotoxic T lymphocyte; boosting; poxvirus vector; PVV; pathogen; tumour;
 KW malaria; parasite; P. falciparum; viral; bacterial; parasitic; cancer;
 KW melanoma; HIV; breast; colon; vaccination.
 XX
 XX Plasmodium falciparum.

OS WO9056919-A2.

PN 17-DEC-1998.

XX 09-JUN-1998; 98WO-GB01681.

XX 09-JUN-1997; 97GB-0011957.

XX (ISIS-) ISIS INNOVATION LTD.

PI Blanchard T, Gilbert SC, Hanke T, Hill AVS, McMichael AJ;

PI Plebanski M, Schneider J, Smith GL;
 XX WPI: 1999-070325/06.
 XX
 XX Generating CD8-positive T cell response to target antigen using
 PT recombinant poxvirus - for treating or preventing malaria and HIV
 PT infection, also epitope strings from Plasmodium and HIV
 XX
 XX Example 1; Page 20; 85pp; English.

XX The invention relates to methods and reagents for generating a protective
 CC CD8+ T-cell immune response against at least one target antigen. The kits
 CC of the invention comprises (i) as priming composition, a source of one or
 CC more CD8+ T-cell (cytotoxic T lymphocytes-CTL) epitopes of the target
 CC antigen, plus a carrier and (ii) as boosting composition a source of CTL
 CC epitopes, with at least one CTL epitope the same as used in (i), with
 CC this source being a non-replicating or replication-impaired recombinant
 CC poxvirus vector (PVV) plus a carrier. If the source of CTL epitopes in
 CC (i) is a viral vector, then the vector in (ii) is from a different virus.
 CC The kits are used to generate an immune response (prophylactic or
 CC therapeutic) against pathogens or tumours, specifically against malaria,
 CC parasites such as P. falciparum, or HIV, and also many other bacterial,
 CC viral or parasitic pathogens. The kits are also used for protective
 CC response against melanoma and cancer of breast or colon, and generally
 CC wherever a strong CD8+ response is protective. The boosting composition
 CC may be used alone to boost a naturally primed response against malaria.
 CC The specified PVV provide an excellent booster effect, better than that
 CC from wild-type poxvirus, resulting in complete rather than partial
 CC protection against sporozoite challenge. Also PVV are safer to use than
 CC wild-type virus.
 XX
 XX Sequence 229 AA;

Query Match 100.0%; Score 44; DB 20; Length 229;
 Best Local Similarity 100.0%; Pred. No. 0.48;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KPNDSLSY 8
 |||||
 Db 29 Kpnksly 36

RESULT 8
 AAY70278
 ID AAY70278 standard; Protein; 350 AA.

XX AC AAY70278;
 XX
 XX 06-JUN-2000 (first entry)
 XX
 XX Recombinant vaccine CDC/NIIMALVAC-1.

XX Recombinant protein; CDC/NIIMALVAC-1; multivalent; malaria; vaccine;
 KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
 KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
 KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
 KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
 KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
 KW Pf27; antiparasitic; prevention; anti-CDC/NIIMALVAC-1 antibody;
 KW honey bee.

XX Chimeric - Apis sp.

OS Chimeric - Clostridium tetani.

OS Chimeric - Plasmodium falciparum.

XX Key Location/Qualifiers

FT Peptide 1..22 /label= Melittin_signal_peptide

FT /note= "Derived from Honey bee"

FT Protein 23..350

FT /label= Mature_CDC/NIIMALVAC-1

FT /note= "Recombinant multivalent malarial vaccine"

```

XX PN W0200011179-A1.
XX PF
XX PD 02-MAR-2000.
XX PF 19-AUG-1999; 99WO-US18869.
XX PR 21-AUG-1998; 98US-0097703.
XX PA (NAIM-) NAT INST IMMUNOLOGY.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Lal AA, Shi YP, Hasnain SE;
XX PF WPI: 2000-237654/20.
XX DR N-PSDB; AAZ51336.
XX
XX Novel recombinant protein as vaccine for treating malarial infection
PT comprises antigenic peptides obtained from different stages of
PT Plasmodium falciparum life cycle -
XX
XX Claim 3; Page 43-44; 52pp; English.
XX
XX The present sequence is that of recombinant protein CDC/NIIMALVAC-1,
CC which is a multivalent, multistage malarial vaccine. The recombinant
CC protein comprises, melittin signal peptide, (His)6 tag, T-cell epitope
CC from tetanus toxoid and 21 antigenic epitopes from circumsporozoite
CC protein (CSP), sporozoite surface protein-2 (SSP-2), liver stage
CC antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1), MSP-2, apical
CC membrane antigen-1 (AMA-1), erythrocyte binding antigen-175 (EBA-175),
CC rhoptry associated protein-1 (RAP-1) and gamete specific antigen, Pfg27.
CC These epitopes were obtained at different stages of the life cycle of
CC Plasmodium falciparum. CDC/NIIMALVAC-1 vaccine has antiparasitic
CC activity and can used for treatment and prevention of malarial
CC infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for detecting
CC P. falciparum in biological samples.
XX
XX Sequence 350 AA;
SQ
Query Match 100.0%; Score 44; DB 21; Length 350;
Best Local Similarity 100.0%; Pred. No. 0.75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 KPNDKSLY 8
Db 184 kpdksly 191
|||||||
RESULTS 9
AAR26944
ID AAR26944 standard; Protein: 493 AA.
XX
AC AAR26944;
XX
DT 08-FEB-1993 (first entry)
XX
DE P.falciparum LSA gene C-terminal region.
XX
KW Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;
KW paludism; liver stage-specific antigen.
XX
OS Plasmodium falciparum.
XX
FH Key Location/Qualifiers
FT Region 13..213
FT /label= repetitive_region
FT Region 214..493
FT /label= non-repetitive_region
XX
PN W09213884-A.
XX
PD 20-AUG-1992.

```

```

XX 05-FEB-1992; 92WO-FR00104.
XX PF
XX PR 05-FEB-1991; 91FR-0001286.
XX PA (INSP ) INST PASTEUR.
XX PI
XX PI Druilhe P, Guerin-Marchand C, Guerinmarchand C;
XX DR WPI: 1992-299985/36.
XX DR N-PSDB; AAQ28119.
XX
XX Polypeptide(s) derived from liver stage of Plasmodium falciparum
PT - for vaccination against, treatment of and diagnosis of malaria
XX
XX Claim 2; Fig 8-10; 81pp; French.
XX
XX The 3' part of the P.falciparum liver-stage specific antigen (LSA)
CC gene codes for a polypeptide sequence which carries a T cell epitope
CC characteristic of a protein produced in hepatocytes infected with
CC P.falciparum. The polypeptide can be used in the preparation of
CC vaccines against malaria.
XX
XX Sequence 493 AA;
SQ
Query Match 100.0%; Score 44; DB 13; Length 493;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 KPNDKSLY 8
Db 434 kpdksly 441
|||||||
RESULTS 10
AAU14727
ID AAU14727 standard; Protein: 514 AA.
XX
AC AAU14727;
XX
DT 24-OCT-2001 (first entry)
XX
DE Novel bone marrow polypeptide #136.
XX
KW Bone marrow; diagnostic; therapeutic; gene therapy; antigenic;
KW haematopoiesis; myeloid; lymph cell disorder; tissue regeneration;
KW wound healing; nutritional supplement; immune disorder;
KW severe combined immunodeficiency; SCID.
XX
OS Homo sapiens.
XX
PN W0200157187-A2.
XX
PD 09-AUG-2001.
XX
DE 05-FEB-2001; 2001WO-US03782.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 30-NOV-2000; 2000US-0250683.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Xue AJ;
XX Ren F, Drmanac RT;
XX
XX WPI: 2001-488875/53.
XX DR N-PSDB; AAS23032.
XX
XX Nucleic acids encoding bone marrow polypeptides, useful in diagnostic
PT and gene therapy -

```

```
XX PS Claim 10; Page 133; 392pp; English.
XX CC AAU14602-AAU14794 represent novel bone marrow polypeptides of the
XX CC invention. The proteins and corresponding coding sequences may be used
XX CC in the prevention, diagnosis and treatment of diseases associated with
XX CC inappropriate bone marrow polypeptide expression. For example, to treat
XX CC disorders associated with decreased expression by rectifying mutations
XX CC or deletions in a patient's genome that affect the activity of the
XX CC polypeptides by expressing inactive proteins or to supplement the
XX CC patient's own production of the polypeptide. Additionally, the nucleic
XX CC acids may be used to produce the polypeptides, by inserting the nucleic
XX CC acids into a host cell and culturing the cell to express the protein.
XX CC The nucleic acid and its complementary sequences may also be used as DNA
XX CC probes in diagnostic assays to detect and quantitate the presence of
XX CC similar nucleic acid sequences in samples, and therefore which patients
XX CC may be in need of restorative therapy. The proteins may also be used as
XX CC antigens in the production of antibodies against bone marrow proteins
XX CC and in assays to identify modulators of their expression and activity.
XX CC The anti-bone marrow protein antibodies and antagonists may also be used
XX CC to down regulate expression and activity. The antibodies may also be used
XX CC as diagnostic agents for detecting the presence of the protein in samples
XX CC (e.g. by enzyme linked immunosorbant assay (ELISA)). The proteins
XX CC may be used to regulate haematopoiesis activity, and consequently in the
XX CC treatment of myeloid or lymph cell disorders; in tissue regeneration,
XX CC such as wound healing; as a nutritional supplement; and in treatment of
XX CC immune disorders such as severe combined immunodeficiency (SCID).
XX SQ Sequence 514 AA;

Query Match 81.8%; Score 36; DB 22; Length 514;
Best Local Similarity 85.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNDKSLY 8
Db IIIIIII
93 pndarsly 99

RESULT 11
AAG28611
ID AAG28611 standard; Protein: 238 AA.
AC AAG28611;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 33894.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127482.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 28-APR-1999; 99US-0130891.
XX PR 30-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 04-MAY-1999; 99US-0132407.
XX PR 05-MAY-1999; 99US-0132484.
XX PR 06-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 07-MAY-1999; 99US-0132487.
XX PR 11-MAY-1999; 99US-0132863.
XX PR 14-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-0134768.
XX PR 19-MAY-1999; 99US-0134941.
XX PR 20-MAY-1999; 99US-0135124.
XX PR 21-MAY-1999; 99US-0135353.
XX PR 21-MAY-1999; 99US-0135629.
XX PR 25-MAY-1999; 99US-0136021.
XX PR 27-MAY-1999; 99US-0136392.
XX PR 28-MAY-1999; 99US-0136782.
XX PR 01-JUN-1999; 99US-0137222.
XX PR 03-JUN-1999; 99US-0137528.
XX PR 04-JUN-1999; 99US-0137502.
XX PR 07-JUN-1999; 99US-0137724.
XX PR 08-JUN-1999; 99US-0138094.
XX PR 10-JUN-1999; 99US-0138540.
XX PR 14-JUN-1999; 99US-0138847.
XX PR 16-JUN-1999; 99US-0139119.
XX PR 16-JUN-1999; 99US-0139452.
XX PR 16-JUN-1999; 99US-0139453.
XX PR 17-JUN-1999; 99US-0139492.
XX PR 18-JUN-1999; 99US-0139454.
XX PR 18-JUN-1999; 99US-0139455.
XX PR 18-JUN-1999; 99US-0139456.
XX PR 18-JUN-1999; 99US-0139457.
XX PR 18-JUN-1999; 99US-0139458.
XX PR 18-JUN-1999; 99US-0139459.
XX PR 18-JUN-1999; 99US-0139460.
XX PR 18-JUN-1999; 99US-0139461.
XX PR 18-JUN-1999; 99US-0139462.
XX PR 18-JUN-1999; 99US-0139463.
XX PR 18-JUN-1999; 99US-0139750.
XX PR 18-JUN-1999; 99US-0139763.
XX PR 21-JUN-1999; 99US-0139817.
XX PR 22-JUN-1999; 99US-0139899.
XX PR 23-JUN-1999; 99US-0140353.
XX PR 23-JUN-1999; 99US-0140354.
XX PR 24-JUN-1999; 99US-0140695.
XX PR 28-JUN-1999; 99US-0140823.
XX PR 29-JUN-1999; 99US-0140991.
XX PR 30-JUN-1999; 99US-0141287.
XX PR 01-JUL-1999; 99US-0141842.
XX PR 01-JUL-1999; 99US-0142154.
XX PR 02-JUL-1999; 99US-0142055.
XX PR 06-JUL-1999; 99US-0142390.
XX PR 08-JUL-1999; 99US-0142803.
XX PR 09-JUL-1999; 99US-0142920.
XX PR 12-JUL-1999; 99US-0142977.
XX PR 13-JUL-1999; 99US-0143542.
XX PR 14-JUL-1999; 99US-0143624.
XX PR 15-JUL-1999; 99US-0144005.
XX PR 16-JUL-1999; 99US-0144085.
XX PR 16-JUL-1999; 99US-0144086.
XX PR 19-JUL-1999; 99US-0144325.
XX PR 19-JUL-1999; 99US-0144331.
XX PR 19-JUL-1999; 99US-0144332.
XX PR 19-JUL-1999; 99US-0144333.
XX PR 19-JUL-1999; 99US-0144334.
XX PR 19-JUL-1999; 99US-0144335.
```

PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 28-JUL-1999; 99US-0145319.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0156917.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 77.3%; Score 34; DB 21; Length 238;
Best Local Similarity 85.7%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNDKSLY 8
Db 140 pndkmly 146

RESULT 12
AAG28610
ID AAG28610 standard; Protein; 260 AA.
XX
AC AAG28610;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 33893.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142303.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145216.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.

PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 77.3%; Score 34; DB 21; Length 260;
Best Local Similarity 85.7%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 PNDKSLY 8
Db 162 pndkmlly 168
RESULT 13
AAG28609
ID AAG28609 standard; Protein; 275 AA.
XX AC AAG28609;
XX DT DT
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 33892.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0128234.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159686.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.
Query Match 77.3%; Score 34; DB 21; Length 275;
Best Local Similarity 85.7%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Caps 0;
QY 2 PNDKSLY 8
Db 177 pndkmy 183
RESULT 14
AAW98800
ID AAW98800 standard; Protein; 380 AA.
XX AC AAW98800;
XX 31-MAR-1999 (first entry)
XX H. pylori GHPO 1287 protein.
XX GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
XX peptic ulcer disease.
XX Helicobacter pylori.
XX WO9843478-A1.
XX 08-OCT-1998.
XX 01-APR-1998; 98WO-US06371.
XX 29-JUL-1997; 97US-0902615.
PR 01-APR-1997; 97US-0833457.
PR 24-JUN-1997; 97US-0881227.
XX (HUMA-) HUMAN GENOME SCI INC.
PA (INNR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;
PI N-PSDB; AAX14519.
DR WPI: 1998-542293/46.
XX N-PSDB; AAX14519.
PT New isolated Helicobacter polynucleotides - used to develop products
PT for the diagnosis, prevention and treatment of Helicobacter
PT infections and gastrointestinal diseases
XX Claim 8; Page 1749-1751; 2054pp; English.
XX This sequence represents a Helicobacter pylori GHPO protein of the
CC invention. The polypeptides can be used for preventing or treating
CC Helicobacter infections, and gastroduodenal diseases associated with
CC these infections, including acute, chronic, and atrophic gastritis, and
CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
CC used for the production of antibodies. The products can also be used for
CC detection and diagnosis.
XX Sequence 380 AA;
SQ

Query Match 77.3%; Score 34; DB 19; Length 380;
Best Local Similarity 75.0%; Pred. No. 77;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Caps 0;

OY 1 KPNDKSLY 8
DB 132 kptckaly 139

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 PNDKSLY 8
DB 93 psckslly 99

Search completed: January 29, 2002, 10:21:45
Job time: 421 sec

RESULT 15

AAB93106
ID AAB93106 standard; Protein: 903 AA.
AC AAB93106;
XX 26-JUN-2001 (first entry)
DT Human protein sequence SEQ ID NO:11968.
DE Human; primer: detection; diagnosis; antisense therapy; gene therapy.
KW Homo sapiens.
OS EPI074617-A2.
PN 07-FEB-2001.
PD 28-JUL-2000; 2000EP-0116126.
PF 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
PA Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
DR WPI: 2001-318749/34.
PT Primer sets for synthesizing polynucleotides, particularly the 5602
full-length cDNAs defined in the specification, and for the detection
and/or diagnosis of the abnormality of the proteins encoded by the
full-length cDNAs -
XX Claim 8; SEQ ID 11968; 2537pp + CD ROM; English.
XX The present invention describes primer sets for synthesizing 5602
full-length cDNAs defined in the specification. Where a primer set
comprises: (a) an oligo-dT primer and an oligonucleotide complementary
to the complementary strand of a polynucleotide which comprises one of
the 5602 nucleotide sequences defined in the specification, where the
oligonucleotide comprises at least 15 nucleotides; or (b) a combination
of an oligonucleotide comprising a sequence complementary to the
complementary strand of a polynucleotide which comprises a 5'-end
sequence and an oligonucleotide comprising a sequence complementary to a
polynucleotide which comprises a 3'-end sequence, where the
the 5'-end sequence/3'-end sequence is selected from those defined in
the specification. The primer sets can be used in antisense therapy and
in gene therapy. The primers are useful for synthesizing polynucleotides,
particularly full-length cDNAs. The primers are also useful for the
detection and/or diagnosis of the abnormality of the proteins encoded by
the full-length cDNAs. The primers allow obtaining of the full-length
cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
represent oligonucleotides, all of which are used in the exemplification
of the present invention.
XX Sequence 903 AA;

Query Match 77.3%; Score 34; DB 22; Length 903;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:24:04 ; Search time 133.18 Seconds
(without alignments)
1.352 Million cell updates/sec

Title: US-09-763-397A-10
Perfect score: 44
Sequence: 1 KPNDKSLY 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	44	100.0	8	2	US-08-318-856A-2	Sequence 2, Appl
2	44	100.0	8	5	PCT-US95-02121-19	Sequence 19, Appl
3	44	100.0	9	2	US-08-318-856A-45	Sequence 45, Appl
4	33	75.0	663	1	US-07-912-015-2	Sequence 2, Appl
5	32	72.7	364	1	US-08-680-726A-56	Sequence 56, Appl
6	32	72.7	364	4	US-09-092-409-56	Sequence 56, Appl
7	32	72.7	967	4	US-09-139-802-201	Sequence 201, Appl
8	31	70.5	45	1	US-07-749-541A-4	Sequence 4, Appl
9	31	70.5	45	1	US-08-275-327-1	Sequence 1, Appl
10	31	70.5	45	5	PCT-US92-03089-1	Sequence 1, Appl
11	30	68.2	345	5	PCT-US95-03323A-2	Sequence 2, Appl
12	30	68.2	378	1	US-08-225-477B-9	Sequence 9, Appl
13	30	68.2	378	5	PCT-US95-04353-9	Sequence 9, Appl
14	30	68.2	656	2	US-08-533-669A-6	Sequence 6, Appl
15	30	68.2	701	4	US-09-132-028-2	Sequence 2, Appl
16	30	68.2	912	5	PCT-US95-03747-2	Sequence 2, Appl
17	29	65.9	6	2	US-08-407-468-44	Sequence 44, Appl
18	29	65.9	10	1	US-07-789-184-45	Sequence 45, Appl
19	29	65.9	10	1	US-07-789-184-60	Sequence 60, Appl
20	29	65.9	10	1	US-07-789-184-66	Sequence 66, Appl
21	29	65.9	10	1	US-07-789-184-67	Sequence 67, Appl
22	29	65.9	10	1	US-07-789-184-203	Sequence 203, Appl
23	29	65.9	10	1	US-08-475-263-45	Sequence 45, Appl
24	29	65.9	10	1	US-08-475-263-60	Sequence 60, Appl
25	29	65.9	10	1	US-08-475-263-66	Sequence 66, Appl
26	29	65.9	10	1	US-08-475-263-67	Sequence 67, Appl
27	29	65.9	10	1	US-08-475-263-203	Sequence 203, Appl

28	29	65.9	10	1	US-08-485-886-45	Sequence 45, Appl
29	29	65.9	10	1	US-08-485-886-60	Sequence 60, Appl
30	29	65.9	10	1	US-08-485-886-66	Sequence 66, Appl
31	29	65.9	10	1	US-08-485-886-67	Sequence 67, Appl
32	29	65.9	10	1	US-08-485-886-203	Sequence 203, Appl
33	29	65.9	10	2	US-08-477-362-45	Sequence 45, Appl
34	29	65.9	10	2	US-08-477-362-60	Sequence 60, Appl
35	29	65.9	10	2	US-08-477-362-66	Sequence 66, Appl
36	29	65.9	10	2	US-08-477-362-67	Sequence 67, Appl
37	29	65.9	10	2	US-08-477-362-203	Sequence 203, Appl
38	29	65.9	10	2	US-08-477-134-45	Sequence 45, Appl
39	29	65.9	10	2	US-08-477-134-60	Sequence 60, Appl
40	29	65.9	10	2	US-08-477-134-66	Sequence 66, Appl
41	29	65.9	10	2	US-08-477-134-67	Sequence 67, Appl
42	29	65.9	10	2	US-08-477-134-203	Sequence 203, Appl
43	29	65.9	10	3	US-08-473-489A-45	Sequence 45, Appl
44	29	65.9	10	3	US-08-473-489A-60	Sequence 60, Appl
45	29	65.9	10	3	US-08-473-489A-66	Sequence 66, Appl

ALIGNMENTS

RESULT 1
US-08-318-856A-2
; Sequence 2, Application US/08318856A
; Patent No. 5972351
; GENERAL INFORMATION:
; APPLICANT: Adrian V.S. Hill, et al.
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM MHC CLASS I-
; TITLE OF INVENTION: RESTRICTED CTL EPITOPES DERIVED FROM PRE-ERYTHROCYTIC STAGE
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,856A
; FILING DATE: October 3, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92 08 068.8
; FILING DATE: April 3, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92 17 704.7
; FILING DATE: August 20, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB93/00711
; FILING DATE: April 5, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: 263-PP1R1577US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 721-8200
; TELEFAX: (202) 721-8250
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum

US-08-318-856A-2

Query Match 100.0%; Score 44; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
|11111111|
DB 1 KPNDKSLY 8

RESULT 2
PCT-US95-02121-19
; Sequence 19, Application PC/TUS9502121
; GENERAL INFORMATION:
; APPLICANT: V.S. Hill, et al.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02121
; FILING DATE: 16-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,484
; FILING DATE: 16-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Patnelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
PCT-US95-02121-19

Query Match 100.0%; Score 44; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNDKSLY 8
|11111111|
DB 1 KPNDKSLY 8

RESULT 3
US-08-318-856A-45
; Sequence 45, Application US/08318856A

Patent No. 5972351
; GENERAL INFORMATION:
; APPLICANT: V.S. Hill, et al.
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM MHC CLASS I-
; TITLE OF INVENTION: RESTRICTED CTL EPITOPES DERIVED FROM PRE-ERYTHROCYTIC STAGE
; TITLE OF INVENTION: ANTIGENS (AS AMENDED)
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,856A
; FILING DATE: October 3, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92 08 068.8
; FILING DATE: April 3, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92 17 704.7
; FILING DATE: August 20, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB93/00711
; FILING DATE: April 5, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: 263-PP1R1577US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 721-8200
; TELEFAX: (202) 721-8250
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-318-856A-45

Query Match 100.0%; Score 44; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
|11111111|
DB 1 KPNDKSLY 8

RESULT 4
US-07-912-015-2
; Sequence 2, Application US/07912015
; Patent No. 5283191
; GENERAL INFORMATION:
; APPLICANT: Morgan, Robin Wilson
; APPLICANT: Willase, Martha Jacoba
; APPLICANT: Claassen, Johannes
; APPLICANT: Sodenmeijer, Paulus Jacobus Antonius
; TITLE OF INVENTION: Marek's Disease Virus vaccine
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Organon Teknika Corporation
; ADDRESSEE: Biotechnology Research Institute
; STREET: 1330-A Piccard Drive
; CITY: Rockville

STATE: Maryland
COUNTRY: U.S.A.
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/912,015
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/615,211
FILING DATE: 16-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Donna Bobrowicz
REGISTRATION NUMBER: 32,196
TELEPHONE: (301)258-5200
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 663 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-912-015-2

Query Match 75.0%; Score 33; DB 1; Length 663;
Best Local Similarity 75.0%; Pred. No. 65;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
||:|:| ||

Db 99 KPDDKLLY 106

RESULT 5
US-08-680-726A-56
; Sequence 56, Application US/08680726A
; Patent No. 5804197
; GENERAL INFORMATION:
; APPLICANT: Haanes, Elizabeth J.
; APPLICANT: Frank, Rexann S.
; TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/680,726A
; FILING DATE: 12-JUL-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-46-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-680-726A-56

Query Match 72.7%; Score 32; DB 1; Length 364;
Best Local Similarity 62.5%; Pred. No. 54;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
||||:|

Db 145 KPNDSGVY 152

RESULT 6
US-09-092-409-56
; Sequence 56, Application US/09092409
; Patent No. 6159478
; GENERAL INFORMATION:
; APPLICANT: Haanes, Elizabeth J.
; APPLICANT: Frank, Rexann S.
; TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/092,409
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/680,726
; FILING DATE: 12-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-46-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-092-409-56

Query Match 72.7%; Score 32; DB 4; Length 364;
Best Local Similarity 62.5%; Pred. No. 54;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
||||:|

Db 145 KPNDSGVY 152

RESULT 7
US-09-139-802-201
; Sequence 201, Application US/09139802
; Patent No. 6180084
; GENERAL INFORMATION:

APPLICANT: Ruoslahti, Erkki
TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing
TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using
TITLE OF INVENTION: Same
FILE REFERENCE: P-LJ 3203
CURRENT APPLICATION NUMBER: US/09/139,802
CURRENT FILING DATE: 1998-08-25
EARLIER APPLICATION NUMBER: 08/926,914
EARLIER FILING DATE: 1997-09-10
EARLIER APPLICATION NUMBER: 08/710,067
EARLIER FILING DATE: 1996-09-10
NUMBER OF SEQ ID NOS: 226
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 201
LENGTH: 967
TYPE: PRT
ORGANISM: Homo sapiens
US-09-139-802-201

Query Match 72.7%; Score 32; DB 4; Length 967;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNDKSLY 8
DB 95 PNDKGLY 101

RESULT 8
US-07-749-541A-4
Sequence 4, Application US/07749541A
Patent No. 5317009
GENERAL INFORMATION:
APPLICANT: Lee-Huang, Sylvia
APPLICANT: Kung, Hsiang-fu
APPLICANT: Huang, Philip L.
APPLICANT: Huang, Paul L.
APPLICANT: Huang, Peter
APPLICANT: Huang, Henry I.
APPLICANT: Chen, Hao-chia
TITLE OF INVENTION: ANTI-HIV PROTEINS, GAP 31, DAP 30 AND DAP 32,
TITLE OF INVENTION: DNA CODING THEREFOR AND THERAPEUTIC USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/749,541A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Livnat, Shmuel
REGISTRATION NUMBER: 33,949
REFERENCE/DOCKET NUMBER: LEE-HUANG-3
TELEPHONE: (202)628-5197
TELEFAX: (202)737-3528
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: TAP 29
US-07-749-541A-4
Query Match 70.5%; Score 31; DB 1; Length 45;
Best Local Similarity 71.4%; Pred. No. 8.3;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 PNDKSLY 8
DB 26 PNEKKLY 32
RESULT 9
US-08-275-327-1
Sequence 1, Application US/08275327
Patent No. 5532214
GENERAL INFORMATION:
APPLICANT: Lee-Huang, Sylvia
APPLICANT: Huang, Philip L.
APPLICANT: Kung, Hsiang-fu
APPLICANT: Huang, Paul L.
APPLICANT: Huang, Peter
APPLICANT: Huang, Henry I.
APPLICANT: Chen, Hao-chia
TITLE OF INVENTION: An Anti-HIV Protein, TAP 29, From
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street, N.W.
CITY: Washington
STATE: DC
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/275,327
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/966,600
FILING DATE:
APPLICATION NUMBER: US/07/685,126
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Livnat, Shmuel
REGISTRATION NUMBER: 33,949
REFERENCE/DOCKET NUMBER: LEE-HUANG-2
TELEPHONE: (202)628-5197
TELEFAX: (202)737-3528
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-275-327-1
Query Match 70.5%; Score 31; DB 1; Length 45;
Best Local Similarity 71.4%; Pred. No. 8.3;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 PNDKSLY 8
DB 26 PNEKKLY 32

Db 26 PNEKKLY 32

RESULT 10

PCT-US92-03089-1

; Sequence 1, Application PC/TUS9203089

; GENERAL INFORMATION:

; APPLICANT: Lee-Huang, Sylvia

; APPLICANT: Huang, Philip L.

; APPLICANT: Kung, Hsiang-fu

; APPLICANT: Huang, Paul L.

; APPLICANT: Huang, Peter

; APPLICANT: Huang, Henry I.

; APPLICANT: Chen, Hao-chia

; TITLE OF INVENTION: An Anti-HIV Protein, TAP 29, From

; TITLE OF INVENTION: Trichosanthes, DNA Coding Therefor and Therapeutic

; TITLE OF INVENTION: Uses Thereof

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Browdy and Neimark

; STREET: 419 Seventh Street, N.W.

; CITY: Washington

; STATE: DC

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/03089

; FILING DATE: 19920415

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Livnat, Shmuel

; REGISTRATION NUMBER: 33,949

; REFERENCE/DOCKET NUMBER: LEE-HUANG-2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)628-5197

; TELEFAX: (202)737-3528

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 45 amino acids

; TYPE: AMINO ACID

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

PCT-US92-03089-1

Query Match 70.5%; Score 31; DB 5; Length 45;

Best Local Similarity 71.4%; Pred. No. 8.3;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNDKSLY 8

11:111

Db 26 PNEKKLY 32

RESULT 11

PCT-US95-03323A-2

; Sequence 2, Application PC/TUS9503323A

; GENERAL INFORMATION:

; APPLICANT: KUNSCH, ET AL.

; TITLE OF INVENTION: Human B-Cell Translocation Genes-2 and 3

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

; ADDRESSEE: CECCHI, STEWART & OULSTEIN

; STREET: 6 BECKER FARM ROAD

; CITY: ROSELAND

; STATE: NEW JERSEY

; COUNTRY: USA

; ZIP: 07068

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 INCH DISKETTE

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WORD PERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/03323A

; FILING DATE: Submitted herewith

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: FERRARO, GREGORY D.

; REGISTRATION NUMBER: 36,134

; REFERENCE/DOCKET NUMBER: 325800-262

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-994-1700

; TELEFAX: 201-994-1744

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 345 AMINO ACIDS

; TYPE: AMINO ACID

; STRANDEDNESS:

; TOPOLOGY: LINEAR

; MOLECULE TYPE: PROTEIN

PCT-US95-03323A-2

Query Match 68.2%; Score 30; DB 5; Length 345;

Best Local Similarity 62.5%; Pred. No. 1.3e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPDKSLY 8

1111111

Db 86 KPDKSVW 93

RESULT 12

US-08-225-477B-9

; Sequence 9, Application US/08225477B

; Patent No. 5635370

; GENERAL INFORMATION:

; APPLICANT: Susan Hockfield

; APPLICANT: Diane M. Jaworski

; TITLE OF INVENTION: BEHAB, A Brain Hya-

; TITLE OF INVENTION: luronan-Binding Protein

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: St. Onge Steward Johnston & Reens

; STREET: 986 Bedford Street

; CITY: Stamford

; STATE: CT

; COUNTRY: United States

; ZIP: 06905

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" 1.44 Mb diskette

; COMPUTER: IBM PC

; OPERATING SYSTEM: MS DOS

; SOFTWARE: Word Processor

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/225,477B

; FILING DATE: April 8, 1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Mary M. Krinsky

; REGISTRATION NUMBER: 32423

; REFERENCE/DOCKET NUMBER: 1751-P0004

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 203-324-6155

; TELEFAX: 203-327-1096

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

LENGTH: 378 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: polypeptide
FRAGMENT TYPE: entire sequence
IMMEDIATE SOURCE: cat cortex
FEATURE:
NAME/KEY: cat brain BEHAB
OTHER INFORMATION: polypeptide encoded by (and set out under) SEQ ID NO 2
US-08-225-477B-9

Query Match 68.2%; Score 30; DB 1; Length 378;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KPNDKSLY 8
:||||:|
Db 128 RPNDSGIY 135

RESULT 13
PCT-US95-04353-9
; Sequence 9, Application PC/TUS9504353
; GENERAL INFORMATION:
; APPLICANT: Susan Hockfield
; APPLICANT: Diane M. Jaworski
; TITLE OF INVENTION: BEHAB, A Brain Hya-
; TITLE OF INVENTION: Luronan-Binding Protein
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: St. Onge Steward Johnston & Reens
; STREET: 986 Bedford Street
; CITY: Stamford
; STATE: CT
; COUNTRY: United States
; ZIP: 06905
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" 1.44 Mb diskette
; COMPUTER: IBM PC
; OPERATING SYSTEM: MS DOS
; SOFTWARE: Word Processor
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04353
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/225,477
; FILING DATE: April 8, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary M. Krinsky
; REGISTRATION NUMBER: 32423
; REFERENCE/DOCKET NUMBER: 1751-P0004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 203-324-6155
; TELEFAX: 203-327-1096
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 378 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: polypeptide
; FRAGMENT TYPE: entire sequence
; IMMEDIATE SOURCE: cat cortex
; FEATURE:
; NAME/KEY: cat brain BEHAB
; OTHER INFORMATION: polypeptide encoded by (and set out under) SEQ ID NO 2
PCT-US95-04353-9

Query Match 68.2%; Score 30; DB 5; Length 378;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KPNDKSLY 8
:||||:|
Db 128 RPNDSGIY 135

RESULT 14
US-08-533-669A-6
; Sequence 6, Application US/08533669A
; Patent No. 5834592
; GENERAL INFORMATION:
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/533,669A
; FILING DATE: 22-SEP-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.420
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 656 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-533-669A-6

Query Match 68.2%; Score 30; DB 2; Length 656;
Best Local Similarity 62.5%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KPNDKSLY 8
:||||:|
Db 412 KPEQKSIY 419

RESULT 15
US-09-132-028-2
; Sequence 2, Application US/09132028
; Patent No. 6222014
; GENERAL INFORMATION:
; APPLICANT: Wilding, Edwina Imogen
; APPLICANT: Black, Michael Trevor
; APPLICANT: Traini, Christopher M.
; TITLE OF INVENTION: nrde
; FILE REFERENCE: GM10162
; CURRENT APPLICATION NUMBER: US/09/132,028
; CURRENT FILING DATE: 1998-08-10

```
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Staphylococcus Aureus
US-09-132-028-2
```

```
Query Match      68.2%   Score 30; DB 4; Length 701;
Best Local Similarity 75.0%   Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 KPNDKSLY 8
   | | | | |
Db 108 KTNDKSOY 115
```

```
Search completed: January 29, 2002, 10:24:04
Job time: 510 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:26:38 ; Search time 144.96 seconds
(without alignments)
4.204 Million cell updates/sec

Title: US-09-763-397A-10

Perfect score: 44

Sequence: 1 KPNKSLY 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_68:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	1909	2 A45592	liver stage antige
2	36	81.8	649	2 T27232	hypothetical prote
3	36	81.8	830	2 T37973	rad16 nucleotide e
4	35	79.5	417	2 S34433	site-specific DNA-
5	35	79.5	504	2 G82631	glutamine syntheta
6	35	79.5	1246	2 T51085	related to protein
7	34	77.3	344	2 T46066	hypothetical prote
8	34	77.3	380	2 B64533	cystathionine gamm
9	34	77.3	380	2 D71973	probable cystathio
10	33	75.0	120	2 T16190	myohemerythrin - p
11	33	75.0	166	2 T25773	hypothetical prote
12	33	75.0	403	2 S42532	hypothetical prote
13	33	75.0	760	2 S55520	chitin synthase (E
14	32	72.7	28	2 B56779	term 5'-region lea
15	32	72.7	237	2 S77151	sugar fermentation
16	32	72.7	240	2 T17583	hypothetical prote
17	32	72.7	405	2 T21321	hypothetical prote
18	32	72.7	406	2 C43664	tryptophan synthas
19	32	72.7	655	2 F72300	oligopeptide ABC t
20	32	72.7	652	1 S57592	probable phosphore
21	32	72.7	728	2 T51817	peroxisomal target
22	32	72.7	967	2 A30325	membrane alanyl am
23	32	72.7	2143	2 G96595	hypothetical prote
24	31	70.5	45	2 A39598	TAP-29 anti-Hiv pr
25	31	70.5	184	2 G70242	conserved hypothet
26	31	70.5	190	2 C70231	conserved hypothet
27	31	70.5	198	2 F45189	chitin synthase (E
28	31	70.5	213	2 S45857	probable membrane
29	31	70.5	230	2 T33622	hypothetical prote

30	31	70.5	260	2 D81191	hypothetical prote
31	31	70.5	443	1 E64667	multidrug-efflux t
32	31	70.5	443	2 F71848	probable transport
33	31	70.5	470	2 S76942	hypothetical prote
34	31	70.5	496	2 S61966	hypothetical prote
35	31	70.5	506	2 A82876	oligopeptide trans
36	31	70.5	553	2 A81057	hypothetical integ
37	31	70.5	569	2 E64427	hypothetical prote
38	31	70.5	612	2 C72293	hypothetical prote
39	31	70.5	780	2 S44560	alpha-alpha-trehal
40	31	70.5	1039	2 G86203	probable N-arginin
41	31	70.5	1355	2 T32092	hypothetical prote
42	30	68.2	134	2 T42915	hypothetical prote
43	30	68.2	186	2 F64693	hypothetical prote
44	30	68.2	188	2 F45188	chitin synthase (E
45	30	68.2	189	2 S55727	chitin synthase (E

ALIGNMENTS

RESULT 1

A45592

liver stage antigen LSA-1 - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 22-Nov-1993 #sequence_revision 02-Dec-1994 #text_change 09-Jun-2000

C:Accession: S24597; A45592; S29393; S34842; B45592; C45592; D45592

R:Zhu, J.; Hollingdale, M.

submitted to the EMBL Data Library, November 1990

A:Reference number: S24597

A:Accession: S24597

A:Molecule type: DNA

A:Residues: 1-1909 <ZHU>

A:Cross-references: EMBL:X56203; NID:g9915; PID:g9916

R:Zhu, J.; Hollingdale, M.R.

Mol. Biochem. Parasitol. 48, 223-226, 1991

A:Title: Structure of Plasmodium falciparum liver stage antigen-1.

A:Reference number: A45592; MUID:92107224

A:Accession: A45592

A:Molecule type: DNA

A:Residues: 1-195:638-688;1165-1215;1590-1909 <ZH2>

A:Note: sequence extracted from NCBI backbone (NCBIN:75010, NCBIN:75012, NCBIN:75014, R:Guerin-Marchand, C.; Drulhe, P.; Galey, B.; Londono, A.; Patatapotikul, J.; Beaudou Nature 329, 164-167, 1987

A:Title: A liver-stage-specific antigen of Plasmodium falciparum characterized by gen

A:Reference number: S29393; MUID:87315391

A:Accession: S29393

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 323-387 <GUE1>

A:Cross-references: EMBL:M28266

R:Guerin-Marchand, C.; Drulhe, P.; Galey, B.; Londono, A.; Patatapotikul, J.; Beaudou submitted to the EMBL Data Library, April 1992

A:Description: a liver-stage-sepcific antigen of plasmodium falciparum characterized

A:Reference number: S34842

A:Accession: S34842

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 323-381; 'HKAI' <GUE2>

A:Cross-references: EMBL:M28266

A:Note: difference at carboxyl end due to frameshift error

C:Comment: This protein is found as flocculent material in the parasitophorous vacuol

C:Superfamily: trichohyalin; calmodulin repeat homology

C:Keywords: EF hand

F:154-1629/region: 17-residue repeats (A-K-E-K-L-Q-E-Q-Q-S-D-L-E-Q-E-R-R)

Query Match 100.0%; Score 44; DB 2; Length 1909;

Best Local Similarity 100.0%; Pred. No. 1;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPNKSLY 8

|||||||

Db 1850 KPNKSLY 1857

RESULT 2

T27232
 Hypothetical protein Y57G11C.20 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T27232
 R:McMurray, A.
 submitted to the EMBL Data Library, September 1997
 A:Reference number: Z20330
 A:Accession: T27232
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-649 <WIL>
 A:Cross-references: EMBL:Z99281; PIDN:CA616521.1; GSPDB:GN00022; CESP:Y57G11C.20
 A:Experimental source: clone Y57G11C
 C:Genetics:
 A:Gene: CESP:Y57G11C.20
 A:Map position: 4

Query Match 81.8%; Score 36; DB 2; Length 649;
 Best Local Similarity 62.5%; Pred. No. 16;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNKSLY 8

|||||:

Db 478 KPNDRNTY 485

RESULT 3

T37973
 rad16 nucleotide excision repair protein homolog - fission yeast (Schizosaccharomyces po
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T37973; T40569
 R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
 submitted to the EMBL Data Library, March 1996
 A:Reference number: Z21759
 A:Accession: T37973
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-830 <WUR>
 A:Cross-references: EMBL:Z69909; PIDN:CAA93784.1; GSPDB:GN00066; SPDB:SPAC19G10.02
 A:Experimental source: strain 972h-; cosmid c19G10
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Davis, P.; Churcher, C.M.
 submitted to the EMBL Data Library, July 1999
 A:Reference number: Z21937
 A:Accession: T40569
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-830 <WOO>
 A:Cross-references: EMBL:AL096788; PIDN:CA846673.1; GSPDB:GN00067; SPDB:SPBC582.10C
 A:Experimental source: strain 972h-; cosmid c582
 C:Genetics: <WUR2>
 A:Gene: SPAC19G10.02
 A:Map position: 1
 A:Introns: 263/3
 C:Genetics: <WOO2>
 A:Gene: SPBC582.10C
 A:Map position: 2

Query Match 81.8%; Score 36; DB 2; Length 830;
 Best Local Similarity 85.7%; Pred. No. 21;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNDKSLY 8

|||||:

Db 638 PNDKTLY 644

RESULT 4

S34433
 site-specific DNA-methyltransferase (adenine-specific) (EC 2.1.1.72) KpnI - Klebsiell
 C:Species: Klebsiella pneumoniae
 A:Variety: ATCC 55014
 C:Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 05-May-2000
 C:Accession: S34433; S34432; S21959
 R:Chatterjee, D.K.; Hammond, A.W.; Blakesley, R.W.; Adams, S.M.; Gerard, G.F.
 submitted to the EMBL Data Library, December 1991
 A:Reference number: S34433
 A:Accession: S34433
 A:Molecule type: DNA
 A:Residues: 1-417 <CHAI>
 A:Cross-references: EMBL:M76435; NID:g149226; PIDN:AAA25090.1; PID:g149228
 R:Chatterjee, D.K.; Hammond, A.W.; Blakesley, R.W.; Adams, S.M.; Gerard, G.F.
 Nucleic Acids Res. 19, 6505-6509, 1991
 A:Title: Genetic organization of the KpnI restriction-modification system.
 A:Reference number: S34431; MUID:92093610
 A:Accession: S34432
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 44-75;93-142;340-359 <CHAZ>
 A:Cross-references: EMBL:M76435
 A:Experimental source: ATCC 55014
 R:Chatterjee, D.
 submitted to the EMBL Data Library, July 1991
 A:Reference number: S21958
 A:Accession: S21959
 A:Molecule type: DNA
 A:Residues: 1-417 <CHA>
 A:Cross-references: EMBL:X61796; NID:g43887; PIDN:CAA43898.1; PID:g43889
 C:Keywords: methyltransferase; restriction modification system; S-adenosylmethionine

Query Match 79.5%; Score 35; DB 2; Length 417;

Best Local Similarity 75.0%; Pred. No. 16;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNKSLY 8

|||||:

Db 182 OPDDKSLY 189

RESULT 5

G82631
 glutamine synthetase XF1842 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
 C:Accession: G82631
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: G82631
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-504 <SIM>
 A:Cross-references: GB:AE004005; GB:AE003849; NID:g9106918; PIDN:AAF84648.1; GSPDB:GN
 R:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carr
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Kleger, J.E.; Kuranae, E.E.; La
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
 M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L

A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1842
C:Superfamily: glutamate--ammonia ligase

Query Match 79.5%; Score 35; DB 2; Length 504;
Best Local Similarity 85.7%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNDKSLY 8
||| ||
Db 427 PNDKDLV 433

RESULT 6

T51085
related to protein kinase PAK1 [imported] - Neurospora crassa
N:Alternate names: protein B2A19.190
C:Species: Neurospora crassa
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T51085
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T51085
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1246 <SCH>
A:Cross-references: EMBL:AL390092; GSPDB:GN00116; NCSP:B2A19.190
A:Experimental source: BAC clone B2A19; strain OR74A
C:Genetics:
A:Gene: NCSP:B2A19.190
A:Map position: 6
A:Introns: 95/2

Query Match 79.5%; Score 35; DB 2; Length 1246;
Best Local Similarity 75.0%; Pred. No. 51;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
||:||||
Db 584 KPDKSLY 591

RESULT 7

T46066
hypothetical protein T18N14.80 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 28-Jul-2000
C:Accession: T46066
R:Deiseny, M.; Berger, C.; Cooke, R.; Grellert, F.; Laudie, M.; Mewes, H.W.; Lemcke, K.; submitted to the Protein Sequence Database, December 1999
A:Reference number: Z23013
A:Accession: T46066
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-344
A:Cross-references: EMBL:AL132968
A:Experimental source: cultivar Columbia; BAC clone T18N14
C:Genetics:
A:Map position: 3
A:Introns: 48/3; 144/1
A:Note: T18N14.80
C:Superfamily: Arabidopsis thaliana hypothetical protein T18N14.80

Query Match 77.3%; Score 34; DB 2; Length 344;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNDKSLY 8
|||| |
Db 162 PNDKMLY 168

RESULT 8

B64533
cystathionine gamma-synthase - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 16-Jul-1999
C:Accession: B64533
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R. Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKee, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467
A:Accession: B64533
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-380 <ROM>
A:Cross-references: GB:AE000532; GB:AE000511; NID:g2313184; PIDN:AAD07176.1; PID:g231 C:Superfamily: O-succinylhomoserine (thiol)-lyase

Query Match 77.3%; Score 34; DB 2; Length 380;
Best Local Similarity 75.0%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
||| |:
Db 132 KPNTKALY 139

RESULT 9

D71973
probable cystathionine gamma-synthase - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 16-Jul-1999
C:Accession: D71973
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D. ; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p A:Reference number: A71800; MUID:99120557
A:Accession: D71973
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-380 <ARN>
A:Cross-references: GB:AE001449; GB:AE001439; NID:g4154604; PIDN:AAD05677.1; PID:g415 A:Experimental source: strain J99
C:Genetics:
A:Gene: metB
C:Superfamily: O-succinylhomoserine (thiol)-lyase

Query Match 77.3%; Score 34; DB 2; Length 380;
Best Local Similarity 75.0%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
||| |:
Db 132 KPNTKALY 139

RESULT 10

S16190
myohemerythrin - polychaete (Nereis diversicolor)
C:Species: Nereis diversicolor (sandworm)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Oct-1999
C:Accession: S16190

R: Takagi, T.; Cox, J.A.
FEBS Lett. 285, 25-27, 1991
A: Title: Primary structure of myohemerythrin from the annelid Nereis diversicolor.
A: Reference number: S16190; MUID: 91293305
A: Accession: S16190
A: Status: preliminary
A: Molecule type: protein
A: Residues: 1-120 <PEB>
C: Superfamily: hemerythrin
C: Keywords: iron, oxygen carrier
E: 25,54,56,75,79,106,113/Binding site: 2Fe-O cluster (His, His, Glu, His, His, Asp)

Query Match 75.0%; Score 33; DB 2; Length 120;
Best Local Similarity 85.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNDKSLY 8
||| ||

DB 94 PNDKLLY 100

RESULT 11
T25773
hypothetical protein F47C10.2 - Caenorhabditis elegans
C: Species: Caenorhabditis elegans
C: Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C: Accession: T25773
R: Minx, P.
submitted to the EMBL Data Library, April 1997
A: Description: The sequence of C. elegans cosmid F47C10.
A: Reference number: Z20084
A: Accession: T25773
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-166 <MIN>
A: Cross-references: EMBL:U97003; PIDN: AAB52269.1; GSPDB: GN00023; CESP: F47C10.2
A: Experimental source: strain Bristol N2; clone F47C10
C: Genetics:
A: Gene: CESP: F47C10.2
A: Map position: 5
A: Introns: 24/3; 75/1

Query Match 75.0%; Score 33; DB 2; Length 166;
Best Local Similarity 71.4%; Pred. No. 17;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNDKSLY 8
||| ||

DB 73 PNDKTVY 79

RESULT 12
S42532
hypothetical protein 1 - Synechococcus sp.
C: Species: Synechococcus sp.
C: Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Mar-1999
C: Accession: S42532
R: Jones, M.C.; Jenkins, J.M.; Smith, A.G.; Howe, C.J.
Plant Mol. Biol. 24, 435-448, 1994
A: Title: Cloning and characterisation of genes for tetrapyrrole biosynthesis from the cyanobacterium Synechococcus sp.
A: Reference number: S42531; MUID: 94169298
A: Accession: S42532
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-403 <JON>
A: Note: the source is designated as Anacystis nidulans R2
C: Superfamily: Synechococcus hypothetical protein 1; tetra tricopeptide repeat homology
E: 283-316/Domain: tetra tricopeptide repeat homology <TN1>
E: 317-350/Domain: tetra tricopeptide repeat homology <TN2>
E: 351-384/Domain: tetra tricopeptide repeat homology <TN3>

Query Match 75.0%; Score 33; DB 2; Length 403;
Best Local Similarity 85.7%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNDKSLY 8
||| |||

DB 348 PNDPSLY 354

RESULT 13
S55520
chitin synthase (EC 2.4.1.16) CSH1 - smut fungus (Ustilago maydis)
C: Species: Ustilago maydis (corn smut)
C: Date: 01-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 20-Aug-1999
C: Accession: S55520
R: Xoconostle-Cazares, B.; Leon-Ramirez, C.; Ruiz-Herrera, J.
submitted to the EMBL Data Library, June 1995
A: Description: Multiple chitin synthetase genes in Ustilago maydis. Cloning and characterization of the chitin synthetase gene from Ustilago maydis.
A: Reference number: S55519
A: Accession: S55520
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-760 <XOC>
A: Cross-references: EMBL: X87748; NID: g861150; PID: g861151
C: Keywords: glycosyltransferase; hexosyltransferase

Query Match 75.0%; Score 33; DB 2; Length 760;
Best Local Similarity 62.5%; Pred. No. 80;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
: |||||

DB 458 RFGDKSIY 465

RESULT 14
B56779
tetC 5'-region leader peptide - Enterococcus faecalis transposon Tn916
C: Species: Enterococcus faecalis
C: Date: 08-Sep-1995 #sequence_revision 08-Sep-1995 #text_change 24-Sep-1999
C: Accession: B56779
R: Su, Y.A.; He, P.; Clewell, D.B.
Antimicrob. Agents Chemother. 36, 769-778, 1992
A: Title: Characterization of the tet(M) determinant of Tn916: evidence for regulation of tet(M) expression by tetC.
A: Reference number: A56779; MUID: 92368175
A: Accession: B56779
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-28 <SUI>
A: Cross-references: GB: U09422; NID: g532533; PIDN: AAB60021.1; PID: g532545
A: Note: sequence extracted from NCBI backbone (NCBIN: 111076, NCBIP: 111079)
C: Superfamily: unassigned leader peptides

Query Match 72.7%; Score 32; DB 2; Length 28;
Best Local Similarity 71.4%; Pred. No. 4.3;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNDKSLY 8
: |||||

DB 12 PSDKSIY 18

RESULT 15
S77151
sugar fermentation stimulation protein - Synechocystis sp. (strain PCC 6803)
N: Alternate names: protein sl12014
C: Species: Synechocystis sp.
A: Variety: PCC 6803
C: Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C: Accession: S77151

R: Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996
A: Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.
A: Reference number: S74322; MUID: 97061201
A: Accession: S77151
A: Status: nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-237 <RAN>
A: Cross-references: EMBL: D90908; GB: AB001339; NID: g1652725; PIDN: BAA17709.1; PID: g165279
A: Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C: Genetics:
A: Gene: sfsA
C: Superfamily: sugar fermentation stimulation protein

Query Match 72.7%; Score 32; DB 2; Length 237;
Best Local Similarity 62.5%; Pred No. 39;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 KPNOKSLY 8
DB 130 KPEKSIY 137

Search completed: January 29, 2002, 10:26:39
Job time: 650 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:13:40 ; Search time 80.65 seconds
(without alignments)
3.637 Million cell updates/sec

Title: us-09-763-397A-10
Perfect score: 44
Sequence: 1 KPNKSLY 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	36	81.8	830	1	YBMA_SCHPO	Q10332	schizosach
2	35	79.5	417	1	MTKL_KLEPN	P25238	klebsiella
3	34	77.3	380	1	METB_HELPJ	Q92mw7	helicobacte
4	34	77.3	380	1	METB_HELPJ	P56069	helicobacte
5	34	77.3	411	1	PPAY_CABEL	Q10944	caenorhabdi
6	33	75.0	120	1	HEMW_NERDI	P22761	nereis dive
7	33	75.0	403	1	YCOA_SYNP7	P42460	synecococc
8	33	75.0	760	1	CHSX_USTWA	Q99126	ustilago ma
9	32	72.7	237	1	SFSA_SYNY3	P73664	synecocyst
10	32	72.7	406	1	TRPB_CAUCR	P12290	caulobacter
11	32	72.7	692	1	MR11_YEAST	P32829	saccharomyc
12	32	72.7	966	1	AMPN_HUMAN	P15144	homo sapien
13	31	70.5	45	1	RIP2_TRIKI	P23029	trichosanth
14	31	70.5	198	1	CHS1_USTWA	P23029	ustilago ma
15	31	70.5	213	1	YBL5_YEAST	P30598	ustilago ma
16	31	70.5	569	1	YA22_METJA	Q58428	methanococc
17	31	70.5	699	1	HS83_ARATH	P51818	arabidopsis
18	31	70.5	780	1	TRBB_YEAST	P35172	saccharomyc
19	30	68.2	188	1	CHS2_EXOJE	P30586	exophiala j
20	30	68.2	189	1	CHS2_PHAEX	P30591	phaeococcom
21	30	68.2	189	1	CHS2_XYLBa	P30604	xylohypha b
22	30	68.2	231	1	Y323_MYCPN	P75322	mycoplasma
23	30	68.2	302	1	YBEJ_ECOLI	P37902	escherichia
24	30	68.2	379	1	YAXO_RHISN	P55707	rhizobium s
25	30	68.2	388	1	OGLE_ERWCA	Q59418	erwinia car
26	30	68.2	396	1	PRRC_ECOLI	P17223	escherichia
27	30	68.2	417	1	PGCB_FELCA	P41725	felis silve
28	30	68.2	499	1	GAG_HVIC4	P05887	human immun
29	30	68.2	499	1	GAG_HVIC4	P12494	human immun
30	30	68.2	521	1	OCLN_CANPA	Q28269	canis famli
31	30	68.2	552	1	DP1B_ECOLI	P77510	escherichia
32	30	68.2	583	1	PD25_MYCTU	Q50586	mycobacteri
33	30	68.2	797	1	VG48_HSVSA	Q01033	herpesvirus

34	30	68.2	803	1	SMI6_YEAST	P09959	saccharomyc
35	30	68.2	830	1	PRLR_COLLI	Q90374	columba liv
36	30	68.2	912	1	PGCB_BOVIN	Q28062	bos taurus
37	30	68.2	966	1	AMPN_FELCA	P79171	felis silve
38	30	68.2	988	1	CHSL_EXODE	P30600	exophiala d
39	30	68.2	1147	1	MYSB_ACACA	P19706	acanthamoeb
40	30	68.2	1479	1	RPOD_WHEAT	Q9xps9	triticum ae
41	30	68.2	1527	1	RPOD_WHEAT	P16025	zea mays (m
42	30	68.2	1861	1	APL_THETU	P38536	t amylopull
43	30	68.2	2181	1	STCA_EMENI	Q12397	emericeella
44	29	65.9	124	1	PROL_PHYPO	P22271	physarum po
45	29	65.9	144	1	PPP5_MOUSE	Q60676	mus musculu

ALIGNMENTS

RESULT 1
YBMA_SCHPO
ID YBMA_SCHPO STANDARD; PRT: 830 AA.
AC Q10332;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 93.9 KDA HELICASE C582.10C IN CHROMOSOME II.
GN SPBC582.10C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Wood V., Rajandream M.A., Barrell B.G., Davis P., Churcher C.M.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL096788; CAB46673.1; -
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR000330; SNF2_N.
DR Pfam; PF00271; helicase_C; 1.
DR Pfam; PF00176; SNF2_N; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICc; 1.
KW Hypothetical protein; Nuclear protein; DNA-binding; Helicase;
KW ATP-binding. 262 269 ATP (POTENTIAL).
FT NP_BIND 262 269
FT SITE 384 387 DEAD BOX.
SQ SEQUENCE 830 AA; 93945 MW; 47086968FDD951CB CRC64;

Query Match 81.8%; Score 36; DB 1; Length 830;
Best Local Similarity 85.7%; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNDKSLY 8
|||||
Db 638 PNDKTLV 644

RESULT 2
MTKL_KLEPN

```

ID MTK1_KLEPN STANDARD; PRT; 417 AA.
AC P25238;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE MODIFICATION METHYLASE KPN1 (EC 2.1.1.72) (ADENINE-SPECIFIC
DE METHYLTRANSFERASE KPN1) (M.KPN1).
GN KPN1M.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OK8;
RX MEDLINE=92093610; PubMed=1754388;
RA Chatterjee D.K., Hammond A.W., Blakesley R.W., Adams S.M.,
RA Gerard G.F.;
RT "Genetic organization of the KpnI restriction-modification system.";
RL Nucleic Acids Res. 19:6505-6509(1991).
CC -!- FUNCTION: THIS METHYLASE RECOGNIZES THE DOUBLE-STRANDED SEQUENCE
CC GGTACC. CAUSES SPECIFIC METHYLATION ON A-4 ON BOTH STRANDS, AND
CC PROTECTS THE DNA FROM CLEAVAGE BY THE KPN1 ENDONUCLEASE.
CC -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA ADENINE =
CC S-ADENOSYL-L-HOMOCYSTEINE + DNA 6-METHYLAMINOPURINE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M76435; AAA25090.1; -
DR EMBL; X61796; CAA43898.1; -
DR PIR; S21959; S21959.
DR PIR; S34433; S34433.
DR REBASE; 3436; M.kpn1.
DR InterPro; IPR002295; D21N6_mtfrase.
DR InterPro; IPR002052; N6_Mtase.
DR Pfam; PF01555; N6_N4_Mtase.
DR PRINTS; PR00506; D21N6MTFRASE.
DR PROSITE; PS00092; N6_MTASE; 1.
KW Transferase; Methyltransferase; Restriction system.
SQ SEQUENCE 417 AA; 47577 MW; 42BFAA6F772347BD CRC64;

```

```

Query Match 79.5%; Score 35; DB 1; Length 417;
Best Local Similarity 75.0%; Pred. No. 8.9;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPNDSLSY 8
DB 182 QPDKSLY 189

```

```

RESULT 3
METB_HELPJ STANDARD; PRT; 380 AA.
AC Q9ZMW7;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CYSTATHIONINE GAMMA-SYNTHASE (EC 4.2.99.9) (CGS) (O-SUCCINYLHOMOSERINE
DE (THIOL)-LYASE).
GN METB OR JHP00098.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -!- CATALYTIC ACTIVITY: O-SUCCINYL-L-HOMOSERINE + L-CYSTEINE =
CC CYSTATHIONINE + SUCCINATE (CAN ALSO USE HYDROGEN SULFIDE AND
CC METHANETHIOL AS SUBSTRATES).
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -!- PATHWAY: SECOND STEP IN METHIONINE BIOSYNTHESIS.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE TRANS-SULFURATION ENZYMES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001449; AAD05677.1; -
DR HSSP; P06721; 1CL2.
DR InterPro; IPR000277; Cys_Met_Meta_PP.
DR Pfam; PF01053; Cys_Met_Meta_PP; 1.
DR PROSITE; PS00869; CYS_MET_METAB_PP; FALSE_NEG.
KW Methionine biosynthesis; Lyase; Pyridoxal phosphate;
KW Complete proteome.
FT BINDING 195 195 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 380 AA; 41223 MW; 358BEAE6C5C802BA CRC64;

```

```

Query Match 77.3%; Score 34; DB 1; Length 380;
Best Local Similarity 75.0%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 KPNDSLSY 8
DB 132 KPNTRALY 139

```

```

RESULT 4
METB_HELPJ STANDARD; PRT; 380 AA.
AC P56069;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CYSTATHIONINE GAMMA-SYNTHASE (EC 4.2.99.9) (CGS) (O-SUCCINYLHOMOSERINE
DE (THIOL)-LYASE).
GN METB OR HP0106.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,

```

RA Venter J.C.:
 RT "The complete genome sequence of the gastric pathogen *Helicobacter pylori*."
 RL Nature 388:539-547(1997).
 CC -!- CATALYTIC ACTIVITY: O-SUCCINYL-L-HOMOSERINE + L-CYSTEINE -
 CC CYSTATHIONINE + SUCCINATE (CAN ALSO USE HYDROGEN SULFIDE AND
 CC METHANETHIOL AS SUBSTRATES).
 CC -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -!- PATHWAY: SECOND STEP IN METHIONINE BIOSYNTHESIS.
 CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -!- SIMILARITY: BELONGS TO THE TRANS-SULFURATION ENZYMES FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000532; AAD07176.1; -
 DR HSSP: P06721; 1CL1.
 DR TIGR: HP0106; -
 DR InterPro: IPR000277; Cys_Met_Meta_PP.
 DR Pfam: PF01053; Cys_Met_Meta_PP; 1.
 DR PROSITE: PS00868; CYS_MET_METAB_PP; 1.
 KW Methionine biosynthesis; Lyase; Pyridoxal phosphate;
 KW Complete proteome.
 FT BINDING 195 195 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 FT SEQUENCE 380 AA; 41118 MW; 58A85489B05257ED CRC64;
 SQ
 Query Match 77.3%; Score 34; DB 1; Length 380;
 Best Local Similarity 75.0%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 KPNDKSLY 8
 DB 132 KPNTRALY 139
 ||| |:
 ||| |:
 RESULT 5
 PPAY_CAEEL STANDARD; PRT; 411 AA.
 AC Q10944;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE PUTATIVE ACID PHOSPHATASE B0361.7 PRECURSOR (EC 3.1.3.2).
 GN B0361.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;.
 RA Du 2;.
 RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U00031; AAA50626.1; -

DR HSSP: P20646; 1RPT.
 DR WormPep: B0361.7; CE00836.
 DR InterPro: IPR000560; His_acid_phosphatse.
 DR Pfam: PF00328; acid_phosphat_2.
 DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hypothetical protein; Signal; Glycoprotein; Hydrolase.
 FT SIGNAL 1 13 POTENTIAL.
 FT CHAIN 14 411 PUTATIVE ACID PHOSPHATASE B0361.7.
 FT ACT_SITE 40 40 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 80 80 BY SIMILARITY.
 FT ACT_SITE 278 278 PROTON DONOR (BY SIMILARITY).
 FT DISULFID 152 363 BY SIMILARITY.
 FT DISULFID 205 302 BY SIMILARITY.
 FT DISULFID 338 342 BY SIMILARITY.
 FT CARBOHYD 104 104 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 210 210 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 218 218 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 312 312 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 323 323 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 411 AA; 47110 MW; BA265D808EC8B11C CRC64;
 Query Match 77.3%; Score 34; DB 1; Length 411;
 Best Local Similarity 75.0%; Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 KPNDKSLY 8
 DB 140 KPTDKELY 147
 ||| |:
 ||| |:
 RESULT 6
 HEMM_NERDI STANDARD; PRT; 120 AA.
 ID HEMM_NERDI
 AC P22761;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE MYOHEMERETHRIN (MHR).
 OS Nereis diversicolor (Sandworm) (Hediste diversicolor).
 OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Aciculata;
 OC Phyllodocta; Nereididae; Nereidae; Nereididae; Nereidae.
 OX NCBI_TaxID=6352;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Muscle;
 RX MEDLINE=91293305; PubMed=2065779;
 RA Takagi T., Cox J.A.;
 RT "Primary structure of myohemerythrin from the annelid Nereis
 RT diversicolor".
 RL FEBS Lett. 285:25-27(1991).
 CC -!- FUNCTION: MYOHEMERETHRIN IS AN OXYGEN-BINDING PROTEIN FOUND IN
 CC THE RETRACTOR MUSCLES OF CERTAIN WORMS. THE OXYGEN-BINDING SITE
 CC CONTAINS TWO IRON ATOMS.
 CC -!- SUBUNIT: MONOMER.
 CC -!- TISSUE SPECIFICITY: MUSCLE.
 CC -!- SIMILARITY: BELONGS TO THE HEMERYTHRIN FAMILY.
 CC PIR: S16190; S16190.
 DR HSSP: P02247; 2MHR.
 DR InterPro: IPR002063; Hemerythrin.
 DR Pfam: PF01814; Hemerythrin; 1.
 DR PRINTS: PR00186; HEMERYTHRIN.
 DR PROSITE: PS00350; HEMERYTHRINS; 1.
 KW Oxygen transport; Muscle protein; Metal-binding; Iron.
 FT METAL 25 25 IRON 1 (BY SIMILARITY).
 FT METAL 54 54 IRON 1 (BY SIMILARITY).
 FT METAL 58 58 IRON 1 AND 2 (BY SIMILARITY).
 FT METAL 75 75 IRON 2 (BY SIMILARITY).
 FT METAL 79 79 IRON 2 (BY SIMILARITY).
 FT METAL 108 108 IRON 2 (BY SIMILARITY).
 FT METAL 113 113 IRON 1 AND 2 (BY SIMILARITY).
 SQ SEQUENCE 120 AA; 13641 MW; A793589D582E8B97 CRC64;

Query Match 75.0%; Score 33; DB 1; Length 120;
 Best Local Similarity 85.7%; Pred. No. 6;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNDKSLY 8
 DB 94 PNDKSLY 100

RESULT 7
 YCOA_SYNP7 STANDARD; PRT; 403 AA.

AC Y42460;
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE HYPOTHETICAL 45.0 KDA PROTEIN IN COBA 5' REGION.
 OS *Synechococcus* sp. (strain PCC 7942) (Anacystis nidulans R2).
 OC Bacteria; Cyanobacteria; Chroococcales; *Synechococcus*.
 OX NCBI_TaxID=1140;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94169298; PubMed=8123787;
 RA Jones M.C., Jenkins J.M., Smith A.G., Howe C.J.;
 RT Cloning and characterisation of genes for tetrapyrrole biosynthesis
 from the cyanobacterium *Anacystis nidulans* R2.;
 RL Plant Mol. Biol. 24:435-448(1994).
 CC -1- SIMILARITY: IN THE N-TERMINAL, STRONG TO H. INFLUENZAE HI0653.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-14 IS THE INITIATOR.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; X70966; CAA50301.1;
 DR InterPro: IPR001173; Glycos_transf_2.
 DR InterPro: IPR001440; TPR.
 DR Pfam: PF00535; Glycos_transf_2; 1.
 DR Pfam: PF00515; TPR; 4.
 DR SMART: SM00028; TPR; 3.
 KW Hypothetical protein; Repeat; TPR repeat.
 FT REPEAT 208 243 TPR 1.
 FT REPEAT 244 282 TPR 2.
 FT REPEAT 283 316 TPR 3.
 FT REPEAT 317 350 TPR 4.
 FT REPEAT 351 387 TPR 5.
 SQ SEQUENCE 403 AA; 44998 MW; 4C0AB39B5498BA13 CRC64;

Query Match 75.0%; Score 33; DB 1; Length 403;
 Best Local Similarity 85.7%; Pred. No. 22;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNDKSLY 8
 DB 348 PNDKSLY 354

RESULT 8
 CHSX_USTMA STANDARD; PRT; 760 AA.
 ID CHSX_USTMA
 AC Q99126;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE CHITIN SYNTHASE 1 (EC 2.4.1.16) (CHITIN-UDP ACETYL-GLUCOSAMINYL

DE TRANSFERASE 1).
 GN CHS1.
 OS Ustilago maydis (Smut fungus).
 OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
 OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
 OX NCBI_TaxID=5270;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=RK32 / A2B3;
 RC MEDLINE=97086517; PubMed=8932711;
 RA Xocoostle-Cazares B., Leon-Ramirez C., Ruiz-Herrera J.;
 RT "Two chitin synthase genes from Ustilago maydis";
 RL Microbiology 142:377-387(1996).
 CC -1- FUNCTION: PLAYS A MAJOR ROLE IN CELL WALL BIOGENESIS.
 CC -1- CATALYTIC ACTIVITY: UDP-N-ACETYL-D-GLUCOSAMINE + [1,4-(N-ACETYL-
 BETA-D-GLUCOSAMINYL)](N) = UDP + [1,4-(N-ACETYL-BETA-D-
 GLUCOSAMINYL)](N+1).
 CC -1- SUBCELLULAR LOCATION: PLASMA MEMBRANE-BOUND.
 CC -1- SIMILARITY: BELONGS TO THE CHITIN SYNTHASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; X87748; CAA61027.1;
 DR InterPro: IPR002923; Chitin_synth.
 DR Pfam: PF01644; Chitin_synth; 1.
 DR ProDom: PD002998; Chitin_synth; 1.
 KW Transferase; Glycosyltransferase; Transmembrane; Cell wall;
 KW Multigene family.
 SQ SEQUENCE 760 AA; 85181 MW; 9377000F57410993 CRC64;

Query Match 75.0%; Score 33; DB 1; Length 760;
 Best Local Similarity 62.5%; Pred. No. 44;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPDKSLY 8
 DB 458 RPDKSLY 465

RESULT 9
 SFSA_SYNY3 STANDARD; PRT; 237 AA.
 ID SFSA_SYNY3
 AC P73664;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE SUGAR FERMENTATION STIMULATION PROTEIN HOMOLOG.
 GN SFSA OR SLL2014.
 OS *Synechocystis* sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; *Synechocystis*.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimpo S., Takeuchi C., Mura T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
Synechocystis sp. strain PCC6803. II. Sequence determination of the
 entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 CC -1- SIMILARITY: BELONGS TO THE SFSA FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D90908; BAA17709.1; -
KW Complete proteome.
SQ SEQUENCE 237 AA; 26847 MW; BECC025E7EB3298A CRC64;

Query Match 72.7%; Score 32; DB 1; Length 237;
Best Local Similarity 62.5%; Pred. No. 20;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 KPNDKSLY 8
|| :||:|
DB 130 KPEKSIY 137

RESULT 10

TRPB_CAUCR
ID TRPB_CAUCR STANDARD; PRT; 406 AA.
AC P12290;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DE TRYPTOPHAN SYNTHASE BETA CHAIN (EC 4.2.1.20).
GN TRPB OR CC3544.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=69394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88115177; PubMed=2828322;
RA Ross C.M., Winkler M.E.;
RT "Structure of the Caulobacter crescentus trpFBA operon.";
RL J. Bacteriol. 170:757-768(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.K.R., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uitterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Shapiro L., Venter J.C., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -!- FUNCTION: THE BETA SUBUNIT IS RESPONSIBLE FOR THE SYNTHESIS OF L-
CC TRYPTOPHAN FROM INDOLE AND L-SERINE.
CC -!- CATALYTIC ACTIVITY: L-SERINE + 1-(INDOL-3-YL)GLYCEROL 3-PHOSPHATE
CC = L-TRYPTOPHAN + GLYCERALDEHYDE 3-PHOSPHATE + H(2)O.
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -!- PATHWAY: LAST (FIFTH) STEP IN BIOSYNTHESIS OF TRYPTOPHAN.
CC -!- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA CHAINS (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE TRPB FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M19129; AAA23057.1; -
DR EMBL; AE006013; AAK25506.1; ALT_INIT.

DR PIR; C43664; C43664.
DR HSSP; P00933; 2TSY.
DR TIGR; CC3544; -
DR InterPro; IPR001926; PALP.
DR InterPro; IPR000993; Trp_synth_beta.
DR Pfam; PF00291; PALP; 1.
DR PROSITE; PS00168; TRP SYNTHASE_BETA; 1.
KW Tryptophan biosynthesis; Pyridoxal phosphate; Lyase;
KW Complete proteome.
FT BINDING 99 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 406 AA; 43491 MW; 83A8161FF477511E CRC64;

Query Match 72.7%; Score 32; DB 1; Length 406;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 KPNDKSLY 8
||||| |
DB 6 KPNDYSAY 13

RESULT 11

MR11_YEAST
ID MR11_YEAST STANDARD; PRT; 692 AA.
AC P32829;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DOUBLE-STRAND BREAK REPAIR PROTEIN MRE11.
GN MRE11 OR YMR224C OR YM9959.06C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE OF 1-643 FROM N.A.
RX MEDLINE=95309669; PubMed=7789757;
RA Johzuka K., Ogawa H.;
RT "Interaction of Mre11 and Rad50: two proteins required for DNA repair
RT and meiosis-specific double-strand break formation in Saccharomycetes
RT cerevisiae.";
RL Genetics 139:1521-1532(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RN Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP FUNCTION.
RX MEDLINE=953511198; PubMed=7625279;
RA Ogawa H., Johzuka K., Nakagawa T., Leem S.H., Hagiwara A.H.;
RT "Functions of the yeast meiotic recombination genes, MRE11 and MRE2.";
RL Adv. Biophys. 31:67-76(1995).
RN [4]
RP FUNCTION.
RX MEDLINE=98315380; PubMed=9651580;
RA Paull T.T., Gellert M.;
RT "The 3-prime to 5-prime exonuclease activity of Mre11 facilitates
RT repair of DNA double-strand breaks.";
RL Mol. Cell 1:969-979(1998).
CC -!- FUNCTION: EXONUCLEASE REQUIRED FOR REPAIR DURING VEGETATIVE GROWTH
CC AND FOR INITIATION OF MEIOTIC RECOMBINATION THROUGH THE FORMATION
CC OF DOUBLE-STRAND BREAKS (DSB). WORK IN COMPLEX WITH RAD50.
CC -!- SIMILARITY: BELONGS TO THE MRE11/RAD32 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

```
CC EMBL: D11463; BAA02017.1; -
DR EMBL: Z49939; CAA90195.1; -
DR PIR: S27428; S27428.
DR SGD: S0004837; MRE11.
DR InterPro: IPR003701; DNA_repair.
DR InterPro: IPR000934; Ser_thr_phosphatase.
DR Pfam: PF02549; DNA_repair.1.
KW Hydrolase; Nuclease; DNA repair; Meiosis.
SQ SEQUENCE 692 AA; 77650 MW; D58433E32E852B73 CRC64;

Query Match 72.7%; Score 32; DB 1; Length 692;
Best Local Similarity 75.0%; Pred. NO. 63;
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

QY 1 KPNKSLY 8
DB 62 KPSKSLY 69

RESULT 12
AMPN_HUMAN STANDARD; PRT; 966 AA.
ID AC PI5144; Q16728;
DC 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE AMINOPEPTIDASE N (EC 3.4.11.2) (MICROSOMAL AMINOPEPTIDASE) (GP150)
DE (MYELOID PLASMA MEMBRANE GLYCOPROTEIN CD13).
GN ANPEP OR PEEN OR APN OR CD13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC TISSUE=Intestine;
RX MEDLINE=89005706; PubMed=2901990;
RA Olsen J., Cowell G.M., Koniushofer E., Danielsen E.M.,
RA Moeller J., Lausgren L., Hansen O.C., Welinder K.G., Engberg J.,
RA Hunziker W., Spiess M., Sjoestrom H., Noren O.;
RT "Complete amino acid sequence of human intestinal aminopeptidase N as
RT deduced from cloned cDNA.";
RL FEBS Lett. 238:307-314(1988).
RN [2]
RX SEQUENCE FROM N.A.
RX MEDLINE=89198086; PubMed=2564851;
RA Look A.T., Ashmun R.A., Shapiro L.H., Peiper S.C.;
RT "Human myeloid plasma membrane glycoprotein CD13 (gp150) is identical
RT to aminopeptidase N.";
RL J. Clin. Invest. 83:1299-1307(1989).
RN [3]
RX SEQUENCE OF 1-14 FROM N.A.
RC TISSUE=Intestinal epithelium;
RX MEDLINE=91268079; PubMed=1675638;
RA Shapiro L.H., Ashmun R.A., Roberts W.M., Look A.T.;
RT "Separate promoters control transcription of the human aminopeptidase
RT N gene in myeloid and intestinal epithelial cells.";
RL J. Biol. Chem. 266:11999-12007(1991).
RN [4]
RX CHARACTERIZATION OF ITS RECEPTOR FUNCTION.
RX MEDLINE=92278436; PubMed=1350662;
RA Yeager C.L., Ashmun R.A., Williams R.K., Cardelino C.B.,
RA Shapiro L.H., Look A.T., Holmes K.V.;
RT "Human aminopeptidase N is a receptor for human coronavirus 229E.";
RL Nature 357:420-422(1992).
RN [5]
RX VARIANTS TYR-241 AND PRO-242.
RX MEDLINE=98112453; PubMed=9452074;
RA Lendekel G., Wex T., Arndt M., Frank K., Franke A., Ansoerge S.;
RT "Identification of point mutations in the aminopeptidase N gene by
RT SSCP analysis and sequencing.";
```

```
RL Hum. Mutat. Suppl. 1:S158-S160(1998).
CC -!- FUNCTION: ALSO SERVES AS A RECEPTOR FOR HUMAN CORONAVIRUS 229E.
CC -!- CATALYTIC ACTIVITY: AMINOACYL-PEPTIDE + H(2)O -> AMINO ACID +
CC OLIGOPEPTIDE.
CC -!- COFACTOR: BINDS ONE ZINC ION.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.
CC -!- PTM: SULFATED (BY SIMILARITY).
CC -!- DISEASE: DEFECTS IN APN SEEM TO BE A CAUSE OF VARIOUS TYPES OF
CC LEUKEMIA OR LYMPHOMA.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M1 (ZINC METALLOPROTEASE);
CC ALSO KNOWN AS THE PEEN SUBFAMILY.
CC -!- DATABASE: NAME=PEEN; NOTE=CD guide CD13 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd13.htm"
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X13276; CAA31640.1; -
DR EMBL: M22324; AAA51719.1; -
DR EMBL: M55522; AAA83399.1; -
DR PIR: A30325; A30325.
DR PIR: S01658; S01658.
DR MEROPS: M01.001; -.
DR MIM: 151530; -.
DR InterPro: IPR001930; Aladiptase.
DR InterPro: IPR000130; Zn_MpPeptidse.
DR Pfam: PF01433; Peptidase_M1; 1.
DR PRINTS: PR00756; ALADIPTASE.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Aminopeptidase; Zinc; Glycoprotein;
KW Sulfation; Transmembrane; Duplication; Signal-anchor; Polymorphism.
FT INIT_MET 0
FT DOMAIN 1 7
FT TRANSMEM 31
FT DOMAIN 32 67
FT DOMAIN 68 966
FT METAL 387 387
FT ACT_SITE 388 388
FT METAL 391 391
FT METAL 410 410
FT ACT_SITE 476 476
FT REPEAT 68 479
FT REPEAT 480 899
FT MOD_RES 175 175
FT MOD_RES 418 418
FT MOD_RES 423 423
FT MOD_RES 912 912
FT CARBOHYD 127 127
FT CARBOHYD 233 233
FT CARBOHYD 264 264
FT CARBOHYD 318 318
FT CARBOHYD 526 526
FT CARBOHYD 624 624
FT CARBOHYD 680 680
FT CARBOHYD 734 734
FT CARBOHYD 817 817
FT VARIANT 241 241
FT VARIANT 242 242
FT VARIANT 535 535
FT CONFLICT 602 602
FT CONFLICT 886 886
FT CONFLICT 109412 MW; C95159DBF1DB469F CRC64;
SQ SEQUENCE 966 AA; 109412 MW; C95159DBF1DB469F CRC64;

Query Match 72.7%; Score 32; DB 1; Length 966;
```


Best Local Similarity 71.4%; Pred. No. 91;
Matches 5; Conservative 1; Mismatches 0; Gaps 0;

QY 2 PNDKSLY 8
Db 94 PNDGRGLY 100

RESULT 13
ID RIP2_TRIKI STANDARD; PRT; 45 AA.
AC P23029;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE RIBOSOME-INACTIVATING PROTEIN TAP-29 (RRNA N-GLYCOSIDASE)
DE (EC 3.2.2.22) (FRAGMENT).
OS Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
RN NCBI_TaxID=3677;
[1]
RP SEQUENCE.
RC TISSUE=Tuberous root;
RX MEDLINE=91319727; PubMed=1713684;
RA Lee-Huang S., Huang P.-L., Kung H.-F., Li B.-Q., Huang P.-L., Huang P.-L.,
RA Huang H.-I., Chen H.-C.;
RT "TAP 29: an anti-human immunodeficiency virus protein from
RT Trichosanthes kirilowii that is nontoxic to intact cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:6570-6574(1991).
CC -1- FUNCTION: CAPABLE OF INHIBITING HIV-1 INFECTION AND REPLICATION.
CC IT INACTIVATES EUKARYOTIC 60S RIBOSOMAL SUBUNITS.
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
CC -1- SIMILARITY: TO OTHER BACTERIAL AND PLANTS RIBOSOME-INACTIVATING
CC PROTEINS. BELONGS TO TYPE 1 RIP.
DR PIR: A39598; A39598.
DR HSSP; P09989; LTCS.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PROSITE; PS00275; SHIGA_RICIN; PARTIAL.
KW Protein synthesis inhibitor; Hydrolase; Toxin.
FT NON_TER 45 45
SQ SEQUENCE 45 AA; 5040 MW; EE1178150C1CF7B7 CRC64;

Query Match 70.5%; Score 31; DB 1; Length 45;
Best Local Similarity 71.4%; Pred. No. 5.3;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNDKSLY 8
Db 26 PNEKKLY 32

RESULT 14
ID CHS1_USTMA STANDARD; PRT; 198 AA.
AC P30598;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE CHITIN SYNTHASE 1 (EC 2.4.1.16) (CHITIN-UDP ACETYL-GLUCOSAMINYL
DE TRANSFERASE 1) (FRAGMENT).
GN CHS1.
OS Ustilago maydis (Smut fungus).
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
RN NCBI_TaxID=5270;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92115692; PubMed=1731323;

RA Bowen A.R., Chen-Wu J.L., Momany M., Young R., Szaniszlo P.J.,
RA Robbins P.W.;
RL "Classification of fungal chitin synthases";
Proc. Natl. Acad. Sci. U.S.A. 89:519-523(1992).
CC -1- FUNCTION: PLAYS A MAJOR ROLE IN CELL WALL BIOGENESIS.
CC -1- CATALYTIC ACTIVITY: UDP-N-ACETYL-D-GLUCOSAMINE + [1,4-(N-ACETYL-
BETA-D-GLUCOSAMINYL)](N) -> UDP + [1,4-(N-ACETYL-BETA-D-
GLUCOSAMINYL)](N+1).
CC -1- SUBCELLULAR LOCATION: PLASMA MEMBRANE-BOUND.
CC -1- SIMILARITY: BELONGS TO THE CHITIN SYNTHASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; M82958; AAA34224.1;
DR PIR: F45189; F45189.
DR InterPro; IPR002923; Chitin_synth.
DR Pfam; PF01644; Chitin_Synth; 1.
DR ProDom; PD002998; Chitin_synth; 1.
KW Transferase; Glycosyltransferase; Transmembrane; Cell wall;
KW Multigene family.
FT NON_TER 1 198
FT NON_TER 198 198
SQ SEQUENCE 198 AA; 22206 MW; A2A399FCAB3EF3C3 CRC64;

Query Match 70.5%; Score 31; DB 1; Length 198;
Best Local Similarity 75.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
Db 155 KPGHKSLY 162

RESULT 15
ID YBL5_YEAST STANDARD; PRT; 213 AA.
AC P38212;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 23.9 KDA PROTEIN IN COQ1-UGA5 INTERGENIC REGION.
GN YBR005W OR YBR0111.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=S288C;
RA Domdey H., Gassenhuber H., Obermaier B., Piravandi E.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO YEAST SSH5.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; Z35874; CAA84941.1;
DR PIR: S45857; S45857.
DR SGD; S0000209; YBR005W.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 40 62 POTENTIAL.

us-09-763-397a-10.rsp

Mon Feb 4 15:23:33 2002

SQ SEQUENCE 213 AA; 23886 MW; 0784FB182C16F1B0 CRC64;
Query Match 70.5%; Score 31; DB 1; Length 213;
Best Local Similarity 71.4%; Pred. NO. 29;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 2 PNDKSLY 8
Db 127 PNDKRAAY 133

Search completed: January 29, 2002, 11:13:41
Job time: 817 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:12:11 ; Search time 285.36 seconds
(without alignments)
4.101 Million cell updates/sec

Title: US-09-763-397a-10
Perfect score: 44
Sequence: 1 KPNDKSLY 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 14627329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	264	5 Q9GTX5	Q9GTX5 plasmodium
2	44	100.0	280	5 Q25843	Q25843 plasmodium
3	44	100.0	280	5 Q25844	Q25844 plasmodium
4	44	100.0	280	5 Q25845	Q25845 plasmodium
5	44	100.0	280	5 Q25846	Q25846 plasmodium
6	44	100.0	280	5 Q25847	Q25847 plasmodium
7	44	100.0	280	5 Q25848	Q25848 plasmodium
8	44	100.0	280	5 Q25849	Q25849 plasmodium
9	44	100.0	280	5 Q25850	Q25850 plasmodium
10	44	100.0	280	5 Q25851	Q25851 plasmodium
11	44	100.0	280	5 Q25852	Q25852 plasmodium
12	44	100.0	280	5 Q25853	Q25853 plasmodium
13	44	100.0	280	5 Q25854	Q25854 plasmodium
14	44	100.0	280	5 Q25855	Q25855 plasmodium
15	44	100.0	280	5 Q25888	Q25888 plasmodium
16	44	100.0	280	5 Q25889	Q25889 plasmodium
17	44	100.0	280	5 Q25900	Q25900 plasmodium
18	44	100.0	280	5 Q27243	Q27243 plasmodium
19	44	100.0	493	5 Q25886	Q25886 plasmodium

20	44	100.0	1909	5 Q25893	Q25893 plasmodium
21	36	81.8	631	4 Q9NZG8	Q9NZG8 homo sapien
22	36	81.8	632	4 Q94937	Q94937 homo sapien
23	36	81.8	649	5 O18244	O18244 caenorhabdi
24	35	79.5	248	3 Q74720	Q74720 debaryomyce
25	35	79.5	504	2 Q9PCD9	Q9PCD9 xylella fas
26	35	79.5	1070	2 Q9EZV7	Q9EZV7 pasteurella
27	35	79.5	1080	2 Q9CM43	Q9CM43 pasteurella
28	35	79.5	1245	3 Q9P3E9	Q9P3E9 neurospora
29	35	79.5	2009	3 Q9P855	Q9P855 gibberella
30	34	77.3	344	10 Q9SCT8	Q9SCT8 arabidopsis
31	34	77.3	811	5 Q9U6M2	Q9U6M2 crithidia f
32	33	75.0	166	5 O01560	O01560 caenorhabdi
33	33	75.0	290	5 Q9VI34	Q9VI34 drosophila
34	33	75.0	361	10 Q9FIX4	Q9FIX4 arabidopsis
35	33	75.0	638	12 Q9IEU9	Q9IEU9 turkey herp
36	33	75.0	663	12 Q9E6F2	Q9E6F2 turkey herp
37	33	75.0	1550	4 Q92547	Q92547 homo sapien
38	33	75.0	1606	3 Q9UW87	Q9UW87 candida alb
39	33	75.0	2024	2 Q9EWA3	Q9EWA3 streptomyce
40	32	72.7	28	2 Q47736	Q47736 enterococu
41	32	72.7	173	2 Q9AI85	Q9AI85 helicobacte
42	32	72.7	173	2 Q9AI78	Q9AI78 helicobacte
43	32	72.7	242	4 Q9NPM1	Q9NPM1 homo sapien
44	32	72.7	244	5 Q9VBX8	Q9VBX8 drosophila
45	32	72.7	259	12 Q91335	Q91335 canine herp

ALIGNMENTS

RESULT 1
Q9GTX5 PRELIMINARY; PRT; 264 AA.
ID Q9GTX5;
AC Q9GTX5;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN-1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_taxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FCCL/HN;
RA Shan 2.X., Yu X.B., Li X.R., Ma C.L.;
RT "Molecular cloning and structure of the 3' terminal of liver stage antigen-1 gene of plasmodium falciparum isolate FCCL/HN.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF246996; AAG12324.1;
FT NON_TER
SQ SEQUENCE 264 AA; 31006 MW; 1714D653E8D9D1D7 CRC64;

Query Match 100.0%; Score 44; DB 5; Length 264;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
DB 205 KENDKSLY 212

RESULT 2
Q25843 PRELIMINARY; PRT; 280 AA.
ID Q25843;
AC Q25843;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBlrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.

OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40884; AAB59230.1; -.
FT NON_TER 1
SQ SEQUENCE 280 AA; 32926 MW; 467080F32FAEAD33 CRC64;

Query Match 100.0%; Score 44; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
Db 221 KPNDKSLY 228

RESULT 3
Q25844
ID Q25844 PRELIMINARY; PRT; 280 AA.
AC Q25844;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40885; AAB59231.1; -.
FT NON_TER 1
SQ SEQUENCE 280 AA; 32940 MW; E9708E3CFFAE9A9CF CRC64;

Query Match 100.0%; Score 44; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
Db 221 KPNDKSLY 228

RESULT 4
Q25845
ID Q25845 PRELIMINARY; PRT; 280 AA.
AC Q25845;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40886; AAB59232.1; -.
FT NON_TER 1
SQ SEQUENCE 280 AA; 32908 MW; FDDE929F34DCA7E9 CRC64;

Query Match 100.0%; Score 44; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
Db 221 KPNDKSLY 228

RESULT 5
Q25846
ID Q25846 PRELIMINARY; PRT; 280 AA.
AC Q25846;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40887; AAC41597.1; -.
FT NON_TER 1
SQ SEQUENCE 280 AA; 32882 MW; 46763641C81AFC33 CRC64;

Query Match 100.0%; Score 44; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
Db 221 KPNDKSLY 228

RESULT 6
Q25847
ID Q25847 PRELIMINARY; PRT; 280 AA.
AC Q25847;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL: L40888; AAC41598.1; -
FT NON_TER 1
SQ SEQUENCE 280 AA; 32944 MW; 467085E23BABC27 CRC64;

Query Match 100.0%; Score 44; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
Db 221 KPNDKSLY 228

RESULT 7
Q25848
ID Q25848 PRELIMINARY; PRT; 280 AA.
AC Q25848;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL: L40889; AAC41599.1; -
FT NON_TER 1
SQ SEQUENCE 280 AA; 32939 MW; E77080385FAA0D61 CRC64;

Query Match 100.0%; Score 44; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
Db 221 KPNDKSLY 228

RESULT 8
Q25849
ID Q25849 PRELIMINARY; PRT; 280 AA.
AC Q25849;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.

RP SEQUENCE FROM N.A.
RC STRAIN=KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL: L40890; AAC41600.1; -
FT NON_TER 1
SQ SEQUENCE 280 AA; 32966 MW; E96D255154DEA9CF CRC64;

Query Match 100.0%; Score 44; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
Db 221 KPNDKSLY 228

RESULT 9
Q25850
ID Q25850 PRELIMINARY; PRT; 280 AA.
AC Q25850;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=KENYAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL: L40891; AAC41601.1; -
FT NON_TER 1
SQ SEQUENCE 280 AA; 32989 MW; E96F812CFFAEBED8 CRC64;

Query Match 100.0%; Score 44; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
Db 221 KPNDKSLY 228

RESULT 10
Q25851
ID Q25851 PRELIMINARY; PRT; 280 AA.
AC Q25851;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=KENYAN;
 RX MEDLINE=96065765; PubMed=7477115;
 RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
 RA Hawley W.A., Collins W.E., Lal A.A.;
 RT "Sequence variations in the non-repetitive regions of the liver stage-
 RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
 RT isolates";
 RL Mol. Biochem. Parasitol. 71:291-294(1995).
 DR EMBL: L40833; AAC41603.1; -;
 FT NON_TER 1
 SQ SEQUENCE 280 AA; 32927 MW; 48706BDFCF40ADD3 CRC64;

Query Match 100.0%; Score 44; DB 5; Length 280;
 Best Local Similarity 100.0%; Pred. No. 0.28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPNDKSLY 8
 |||||
 Db 221 KPNDKSLY 228

RESULT 11
 Q25852
 ID Q25852 PRELIMINARY; PRT; 280 AA.
 AC Q25852;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
 GN LSA-1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KENYAN;
 RX MEDLINE=96065765; PubMed=7477115;
 RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
 RA Hawley W.A., Collins W.E., Lal A.A.;
 RT "Sequence variations in the non-repetitive regions of the liver stage-
 RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
 RT isolates";
 RL Mol. Biochem. Parasitol. 71:291-294(1995).
 DR EMBL: L40834; AAB59233.1; -;
 FT NON_TER 1
 SQ SEQUENCE 280 AA; 32927 MW; E6648F85FBAA245 CRC64;

Query Match 100.0%; Score 44; DB 5; Length 280;
 Best Local Similarity 100.0%; Pred. No. 0.28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPNDKSLY 8
 |||||
 Db 221 KPNDKSLY 228

RESULT 12
 Q25853
 ID Q25853 PRELIMINARY; PRT; 280 AA.
 AC Q25853;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
 GN LSA-1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KENYAN;

RX MEDLINE=96065765; PubMed=7477115;
 RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
 RA Hawley W.A., Collins W.E., Lal A.A.;
 RT "Sequence variations in the non-repetitive regions of the liver stage-
 RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
 RT isolates";
 RL Mol. Biochem. Parasitol. 71:291-294(1995).
 DR EMBL: L40835; AAB59234.1; -;
 FT NON_TER 1
 SQ SEQUENCE 280 AA; 32938 MW; 46751C45F4DCBD33 CRC64;

Query Match 100.0%; Score 44; DB 5; Length 280;
 Best Local Similarity 100.0%; Pred. No. 0.28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPNDKSLY 8
 |||||
 Db 221 KPNDKSLY 228

RESULT 13
 Q25854
 ID Q25854 PRELIMINARY; PRT; 280 AA.
 AC Q25854;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
 GN LSA-1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KENYAN;
 RX MEDLINE=96065765; PubMed=7477115;
 RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
 RA Hawley W.A., Collins W.E., Lal A.A.;
 RT "Sequence variations in the non-repetitive regions of the liver stage-
 RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
 RT isolates";
 RL Mol. Biochem. Parasitol. 71:291-294(1995).
 DR EMBL: L40836; AAC41604.1; -;
 FT NON_TER 1
 SQ SEQUENCE 280 AA; 32966 MW; 61851C5AD6E1A211 CRC64;

Query Match 100.0%; Score 44; DB 5; Length 280;
 Best Local Similarity 100.0%; Pred. No. 0.28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPNDKSLY 8
 |||||
 Db 221 KPNDKSLY 228

RESULT 14
 Q25855
 ID Q25855 PRELIMINARY; PRT; 280 AA.
 AC Q25855;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
 GN LSA-1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96065765; PubMed=7477115;
 RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,

RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates.";
RL MOI. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40837; AAC41605.1; -
FT NON_TER 1
SQ SEQUENCE 280 AA; 32866 MW; 73164A709789DD31 CRC64;

Query Match 100.0%; Score 44; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
DB 221 KPNDKSLY 228

RESULT 15
Q25888
ID Q25888 PRELIMINARY; PRT; 280 AA.
AC Q25888;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE LIVER STAGE-SPECIFIC ANTIGEN 1 (FRAGMENT).
GN LSA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRAZILIAN;
RX MEDLINE=96065765; PubMed=7477115;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates.";
RL MOI. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40908; AAB59228.1; -
FT NON_TER 1
SQ SEQUENCE 280 AA; 32936 MW; FDF9A29104C5A7E9 CRC64;

Query Match 100.0%; Score 44; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8
DB 221 KPNDKSLY 228

Search completed: January 29, 2002, 11:12:12
Job time: 768 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:21:45 ; Search time 310.82 seconds
(without alignments)
4.290 Million cell updates/sec

Title: US-09-763-397A-11

Perfect score: 105

Sequence: 1 NSGCFRLDERECKLL 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_1101.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	100.0	18	21	AA70287
2	105	100.0	49	14	AA41354
3	105	100.0	49	14	AA41355
4	105	100.0	96	22	AA37608
5	105	100.0	108	22	AA37609
6	105	100.0	116	18	AA36103
7	105	100.0	116	18	AA22592
8	105	100.0	127	18	AA22593
9	105	100.0	127	18	AA36102
10	105	100.0	350	21	AA70278
11	105	100.0	355	20	AA09372

12	105	100.0	355	20	AA05832	Merozoite surface
13	105	100.0	361	20	AA09373	Merozoite surface
14	105	100.0	361	20	AA05833	Merozoite surface
15	105	100.0	375	22	AA88326	A major merozoite
16	105	100.0	376	20	AA09374	Modified merozoite
17	105	100.0	376	20	AA05834	Modified merozoite
18	105	100.0	1639	19	AAW54145	P. falciparum synt
19	100	95.2	1654	6	AA050777	Sequence of the P1
20	60	57.1	93	22	AA837611	Merozoite surface
21	60	57.1	206	22	AA006669	Plasmodium vivax m
22	53	50.5	106	14	AA41358	P. yoelli combined
23	50	47.6	29	22	AA020558	Peptide #6992 enco
24	50	47.6	29	22	AA035390	Peptide #9427 enco
25	48	45.7	1186	20	AA05839	Banana ripening fr
26	46	43.8	569	20	AA07045	Breast cancer asso
27	45	42.9	75	22	AA18846	Peptide #5280 enco
28	45	42.9	75	22	AA031376	Peptide #5413 enco
29	45	42.9	79	22	AA018850	Peptide #5284 enco
30	45	42.9	79	22	AA031381	Peptide #5418 enco
31	45	42.9	152	21	AA016357	Pinus radiata coum
32	45	42.9	496	21	AA041051	Human OREX ORF815
33	45	42.9	2485	15	AA059921	RAS associated GAP
34	45	42.9	2485	15	AA059922	RAS associated GAP
35	45	42.9	2818	13	AA022268	Nf1 gene product.
36	45	42.9	2818	18	AA013280	Human neurofibroma
37	45	42.9	2818	22	AA05486	Human neurofibroma
38	44	41.9	27	18	AA035723	Kappa-conotoxin pv
39	44	41.9	567	21	AA054838	Human diacylglycer
40	44	41.9	567	21	AA018664	Human diacylglycer
41	43	41.0	236	20	AA04707	Mouse RPTPa amino
42	43	41.0	236	20	AA04700	Human RPTPa amino
43	43	41.0	254	22	AA059372	Human protein tyro
44	43	41.0	454	21	AA091625	Human secreted pro
45	43	41.0	793	13	AA020743	Murine receptor ty

ALIGNMENTS

RESULT 1

AA70287
ID AAY70287 standard; peptide; 18 AA.

XX AC AAY70287;

XX DT 06-JUN-2000 (first entry)

XX DE Plasmodium falciparum MSP-1 antigenic epitope, P597.

XX KW Recombinant protein; CDC/NIAIDVAC-1; multivalent; malaria; vaccine;
KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
KW Pf927; antiparasitic; prevention; anti-CDC/NIAIDVAC-1 antibody.

XX OS Plasmodium falciparum.

XX PN WO200011179-A1.

XX PD 02-MAR-2000.

XX PF 19-AUG-1999; 99WO-US18869.

XX PR 21-AUG-1998; 98US-0097703.

XX PA (NAM-) NAT INST IMMUNOLOGY.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Lal AA, Shi YP, Hasnain SE;

XX WPI; 2000-237654/20.

XX Novel recombinant protein as vaccine for treating malarial infection
 PT comprises antigenic peptides obtained from different stages of
 PT plasmodium falciparum life cycle -
 XX
 PS Claim 2; Page 16; 52pp; English.
 XX The present sequence is the antigenic epitope P597, derived from
 CC merozoite surface protein-1 (MSP-1) of the asexual blood stage of
 CC Plasmodium falciparum. It is used in the construction of recombinant
 CC protein CDC/NIMALVAC-1, which is a multivalent, multistage malarial
 CC vaccine. The recombinant protein comprises, melittin signal peptide,
 CC (his)6 tag, T-cell epitope from tetanus toxoid and 21 antigenic epitopes
 CC from circumsporozoite protein (CSP), sporozoite surface protein-2
 CC (SSP-2), liver stage antigen-1 (LSA-1), merozoite surface protein-1
 CC (MSP-1), apical membrane antigen-1 (AMA-1), erythrocyte binding
 CC antigen-175 (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete
 CC specific antigen, Pf27. These epitopes were obtained at different stages
 CC of the life cycle of P. falciparum. CDC/NIMALVAC-1 vaccine has
 CC antiparasitic activity and can be used for treatment and prevention of
 CC malarial infections. Anti-CDC/NIMALVAC-1 antibodies can be used for
 CC detecting P. falciparum in biological samples.
 XX Sequence 18 AA;

Query Match 100.0%; Score 105; DB 21; Length 18;
 Best Local Similarity 100.0%; Pred. No. 2.8e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
 DB 1 nsgcfhrldereckcll 18

RESULT 2

AAAR41354
 ID AAR41354 standard; peptide; 49 AA.
 AC AAR41354;
 DT 04-MAR-1994 (first entry)
 DE MSP1EGF1A EGFI-like domain variant.
 XX Epidermal growth factor 1; merozoite surface protein 1; malaria;
 KW vaccine.
 XX Plasmodium yoelii.
 OS
 FH Key Location/Qualifiers
 FT Cleavage-site 1 /note= "introduced to facilitate cleavage
 FT from recombinant protein"
 XX WO9317107-A.
 XX 02-SEP-1993.
 XX 22-FEB-1993; 93WO-GB00367.
 XX 22-FEB-1992; 92GB-0003821.
 XX (MEDI-) MEDICAL RES COUNCIL.
 XX Blackman MJ, Chappel JA, Holder AA;
 XX WPI; 1993-288413/36.
 XX Allelic variants of epidermal growth factor 1- or 2-like domains - of
 PT merozoite surface protein 1, produced recombinantly for malaria
 PT vaccines
 XX

PS Claim 1; Fig 1a; 35pp; English.
 XX The sequence is that of an allelic variant of a merozoite surface
 CC protein-1 epidermal growth factor (EGF) 1 like domain. It may be
 CC used alone or as part of a fusion protein of EGF-1-like and
 CC EGF-2-like domains in vaccines against malaria. When expressed
 CC recombinantly it is produced in a form indistinguishable from
 CC that in the native protein.
 XX Sequence 49 AA;

Query Match 100.0%; Score 105; DB 14; Length 49;
 Best Local Similarity 100.0%; Pred. No. 7.2e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
 DB 16 nsgcfhrldereckcll 33

RESULT 3

AAAR41355
 ID AAR41355 standard; peptide; 49 AA.
 AC AAR41355;
 DT 04-MAR-1994 (first entry)
 DE MSP1EGF1B EGFI-like domain variant.
 XX Epidermal growth factor 1; merozoite surface protein 1; malaria;
 KW vaccine.
 XX Plasmodium yoelii.
 OS
 FH Key Location/Qualifiers
 FT Cleavage-site 1 /note= "introduced to facilitate cleavage
 FT from recombinant protein"
 XX WO9317107-A.
 XX 02-SEP-1993.
 XX 22-FEB-1993; 93WO-GB00367.
 XX 22-FEB-1992; 92GB-0003821.
 XX (MEDI-) MEDICAL RES COUNCIL.
 XX Blackman MJ, Chappel JA, Holder AA;
 XX WPI; 1993-288413/36.
 XX Allelic variants of epidermal growth factor 1- or 2-like domains - of
 PT merozoite surface protein 1, produced recombinantly for malaria
 PT vaccines
 XX
 PS Claim 1; Fig 1b; 35pp; English.
 XX The sequence is that of an allelic variant of a merozoite surface
 CC protein-1 epidermal growth factor (EGF) 1 like domain. It may be
 CC used alone or as part of a fusion protein of EGF-1-like and
 CC EGF-2-like domains in vaccines against malaria. When expressed
 CC recombinantly it is produced in a form indistinguishable from
 CC that in the native protein.
 XX Sequence 49 AA;

Query Match 100.0%; Score 105; DB 14; Length 49;
 Best Local Similarity 100.0%; Pred. No. 7.2e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18
 |||
 DB 16 nsgcfhrldereckcll 33

RESULT 4

AAAB37608
 ID AAB37608 standard; protein; 96 AA.

XX
 AC AAB37608;

DT 27-FEB-2001 (first entry)

XX Merozoite surface protein-1.

XX Merozoite surface protein; protazoacide; vaccine; malaria.

OS Plasmodium falciparum.

XX WO200063245-A2.

PD 26-OCT-2000.

XX 20-APR-2000; 2000WO-GB01558.

XX 20-APR-1999; 99GB-0009072.

PR 13-MAY-1999; 99US-0311817.

XX 25-MAY-1999; 99CA-2271451.

XX (MEDI-) MEDICAL RES COUNCIL.

PI Holder A, Birdsall B, Feeney J, Morgan W, Syed S, Uthaipibull C;

XX WPI; 2001-015762/02.

XX Novel variants of the C-terminal fragment of Plasmodium merozoite

PT surface protein-1, useful as vaccines for treating or preventing

PT malaria

XX Example 2; Page 48; 126pp; English.

XX The present invention relates to non-natural variants of a C-terminal
 CC fragment of a Plasmodium merozoite surface protein-1 (MSP-1). The
 CC non-natural variants have reduced affinity for at least 1 antibody
 CC capable of blocking a second antibody that inhibits the proteolytic
 CC cleavage of Plasmodium MSP-1.4.2, and has the same affinity for at least
 CC one third antibody that inhibits the proteolytic cleavage of Plasmodium
 CC MSP-1.4.2, compared to natural MSP-1.1.9. The present sequence is the
 CC wild-type MSP-1 protein. This sequence was used to generate the variants
 CC of the present invention. The non-natural variants of the present
 CC invention are useful for immunising a mammal against malaria, and can be
 CC used to treat malaria.

XX Sequence 96 AA;

Query Match 100.0%; Score 105; DB 22; Length 96;

Best Local Similarity 100.0%; Pred. No. 1.3e-07;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18

DB 15 nsgcfhrldereckcll 32

RESULT 5

AAAB37609
 ID AAB37609 standard; protein; 108 AA.

XX
 AC AAB37609;

XX

DT 27-FEB-2001 (first entry)

XX Merozoite surface protein-119.

KW Merozoite surface protein; protazoacide; vaccine; malaria.

XX Plasmodium falciparum.

XX WO200063245-A2.

PD 26-OCT-2000.

XX 20-APR-2000; 2000WO-GB01558.

XX 20-APR-1999; 99GB-0009072.

PR 13-MAY-1999; 99US-0311817.

XX 25-MAY-1999; 99CA-2271451.

XX (MEDI-) MEDICAL RES COUNCIL.

PI Holder A, Birdsall B, Feeney J, Morgan W, Syed S, Uthaipibull C;

XX WPI; 2001-015762/02.

DR N-PSDB; AAC68977.

XX Novel variants of the C-terminal fragment of Plasmodium merozoite

PT surface protein-1, useful as vaccines for treating or preventing

PT malaria

XX Example 5; Fig 15; 126pp; English.

XX The present invention relates to non-natural variants of a C-terminal
 CC fragment of a Plasmodium merozoite surface protein-1 (MSP-1). The
 CC non-natural variants have reduced affinity for at least 1 antibody
 CC capable of blocking a second antibody that inhibits the proteolytic
 CC cleavage of Plasmodium MSP-1.4.2, and has the same affinity for at least
 CC one third antibody that inhibits the proteolytic cleavage of Plasmodium
 CC MSP-1.4.2, compared to natural MSP-1.1.9. The non-natural variants of the
 CC present invention are useful for immunising a mammal against malaria, and
 CC can be used to treat malaria. The present sequence is MSP-119 protein.

XX Sequence 108 AA;

Query Match 100.0%; Score 105; DB 22; Length 108;

Best Local Similarity 100.0%; Pred. No. 1.5e-07; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18

DB 27 nsgcfhrldereckcll 44

RESULT 6

AAW36103
 ID AAW36103 standard; Protein; 116 AA.

XX
 AC AAW36103;

DT 25-MAR-1998 (first entry)

XX PfMSP1(p19)A protein sequence.

XX Plasmodium vivax; merozoite surface protein; MSP1; p19;

KW Plasmodium falciparum; malaria; vaccine; immunity; epitope.

XX Plasmodium falciparum.

OS Synthetic.

XX Key Location/Qualifiers

XX 1..95

FT Region /note= "amino acids derived from P. falciparum MSP1 p19
 FT fragment"

FT Region 96..116
 XX /note= "glycosylphosphatidylinositol anchoring sequence"
 PF WO9730158-A2.
 XX
 XX
 XX
 XX 21-AUG-1997.
 XX
 XX 14-FEB-1997; 97WO-FR00290.
 XX
 XX 14-FEB-1996; 96FR-0001822.
 XX
 XX (INSP) INST PASTEUR.
 XX (UINY) UNIV NEW YORK STATE.
 XX
 XX Barnwell JW, Longacre-Andre S, Mendis K, Nato F;
 XX Roth C;
 XX WPI; 1997-425033/39.
 XX DR N-PSDB; AAT94550.
 XX
 XX Recombinant protein containing the merozoite surface protein-1 p19
 XX fragment - useful in anti-malarial vaccines, diagnosis and protein
 XX purification
 XX
 XX Disclosure; Fig 1B; 85pp; French.
 XX
 XX This is the amino acid sequence of a recombinant protein comprising
 XX amino acids 1613-1705 of the Plasmodium falciparum merozoite surface
 XX protein 1 (MSP1) 19 kD C-terminal fragment (p19), linked to a
 XX glycosylphosphatidylinositol membrane anchoring sequence. p19 is the
 XX C-terminal fragment of the 42 kD MSP1 from Plasmodium species.
 XX The recombinant protein can be used for the production of anti-malarial
 XX vaccines, where the p19 fragment provides a high level of protective
 XX immunity since it includes epitopes not presented in the p42 fragment.
 XX
 XX Sequence 116 AA;
 SQ
 Query Match 100.0%; Score 105; DB 18; Length 116;
 Best Local Similarity 100.0%; Pred. No. 1.6e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NSGCFRHLDERECKLL 18
 DB 17 NSGCFRHLDERECKLL 34
 RESULT 7
 AAW22592
 ID AAW22592 standard; Protein; 116 AA.
 XX
 XX AAW22592;
 XX
 XX 25-MAR-1998 (first entry)
 XX
 XX PfMSP1(p19)A protein sequence.
 XX
 XX Plasmodium vivax; merozoite surface protein; MSP1; p19;
 XX Plasmodium falciparum; malaria; vaccine; immunity; epitope.
 XX
 XX Plasmodium falciparum.
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 XX Region 1..95
 XX /note= "amino acids derived from P. falciparum MSP1 p19
 XX fragment"
 XX Region 96..116
 XX /note= "glycosylphosphatidylinositol anchoring sequence"
 XX
 XX WO9730159-A2.
 XX
 XX 21-AUG-1997.

XX 14-FEB-1997; 97WO-FR00291.
 XX
 XX 14-FEB-1996; 96FR-0001821.
 XX
 XX (INSP) INST PASTEUR.
 XX (UINY) UNIV NEW YORK STATE.
 XX
 XX Barnwell JW, Longacre-Andre S, Mendis K, Nato F;
 XX Roth C;
 XX WPI; 1997-425034/39.
 XX DR P-PSDB; AAW22592.
 XX
 XX Recombinant protein containing Plasmodium merozoite surface
 XX protein-1 p42 fragment - useful in antimalarial vaccines, also new
 XX antibodies for diagnosis and protein purification
 XX
 XX Disclosure; Fig 1B; 85pp; French.
 XX
 XX This is the amino acid sequence of a recombinant protein comprising
 XX amino acids 1613-1705 of the Plasmodium falciparum merozoite surface
 XX protein 1 (MSP1) 19 kD C-terminal fragment (p19), linked to a
 XX glycosylphosphatidylinositol membrane anchoring sequence. p19 is the
 XX C-terminal fragment of the 42 kD MSP1 from Plasmodium species.
 XX The recombinant protein can be used for the production of anti-malarial
 XX vaccines, where the p19 fragment provides a high level of protective
 XX immunity since it includes epitopes not presented in the p42 fragment.
 XX
 XX Sequence 116 AA;
 SQ
 Query Match 100.0%; Score 105; DB 18; Length 116;
 Best Local Similarity 100.0%; Pred. No. 1.6e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NSGCFRHLDERECKLL 18
 DB 17 NSGCFRHLDERECKLL 34
 RESULT 8
 AAW22593
 ID AAW22593 standard; Protein; 127 AA.
 XX
 XX AAW22593;
 XX
 XX 25-MAR-1998 (first entry)
 XX
 XX PfMSP1(p19)S protein sequence.
 XX
 XX Chimeric; Plasmodium vivax; merozoite surface protein; MSP1; p19;
 XX Plasmodium falciparum; malaria; vaccine; immunity; epitope.
 XX
 XX Chimeric - Plasmodium vivax.
 XX Chimeric - Plasmodium falciparum.
 XX
 XX Key Location/Qualifiers
 XX Peptide 1..19
 XX /note= "signal peptide"
 XX Protein 20..127
 XX /note= "mature protein"
 XX Region 1..32
 XX /note= "derived from P. vivax MSP1"
 XX Region 33..34
 XX /note= "encoded by restriction enzyme sequence used to
 XX create the chimeric sequence"
 XX Region 35..127
 XX /note= "derived from P. falciparum C-terminal p19
 XX fragment of MSP1"
 XX
 XX WO9730159-A2.

```

PD 21-AUG-1997.
XX
PF 14-FEB-1997; 97WO-FR00291.
XX
PR 14-FEB-1996; 96FR-0001821.
XX
PA (INSP ) INST PASTEUR.
XX
PA (UNYK ) UNIV NEW YORK STATE.
XX
PI Barnwell JW, Longacre-Andre S, Mendis K, Nato F;
PI Roth C;
XX
DR WPI: 1997-425034/39.
DR P-PSDB; AAW22592.
XX
XX Recombinant protein containing Plasmodium merozoite surface
PT protein-1 p42 fragment - useful in antimalarial vaccines, also new
PT antibodies for diagnosis and protein purification
XX
XX Disclosure; Fig 1C; 85pp; French.
XX
CC This is the amino acid sequence of a chimeric protein comprising amino
CC acids 1-32 of the Plasmodium vivax merozoite surface protein 1 (MSp1)
CC linked to amino acids Asn1613-Ser1705 which correspond to the 19 kD
CC C-terminal fragment (p19) of MSP-1 from Plasmodium falciparum.
CC p19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium
CC species.
CC The recombinant protein can be used for the production of anti-malarial
CC vaccines, where the p19 fragment provides a high level of protective
CC immunity since it includes epitopes not presented in the p42 fragment.
XX
SQ Sequence 127 AA;

Query Match 100.0%; Score 105; DB 18; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
Db |||||
49 nsqcfrrhldereckcll 66

RESULT 9
ID AAW36102 standard; Protein; 127 AA.
XX
AC AAW36102;
XX
DT 25-MAR-1998 (first entry)
XX
DE
XX
KW Chimeric; Plasmodium vivax; merozoite surface protein; MSP1; p19;
KW Plasmodium falciparum; malaria; vaccine; immunity; epitope.
XX
OS Chimeric - Plasmodium vivax.
OS Chimeric - Plasmodium falciparum.
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT Protein /note= "signal peptide"
FT 20..127
FT Region /note= "mature protein"
FT 1..32
FT Region /note= "derived from P. vivax MSP1"
FT 33..34
FT Region /note= "encoded by restriction enzyme sequence used to
FT create the chimeric sequence"
FT 35..127
FT /note= "derived from P. falciparum C-terminal p19
FT fragment of MSP1"
XX
WN09730158-A2.

```

```

XX 21-AUG-1997.
PD
XX
XX 14-FEB-1997; 97WO-FR00290.
PF
XX
XX 14-FEB-1996; 96FR-0001822.
PR
XX
XX (INSP ) INST PASTEUR.
PA
XX (UNYK ) UNIV NEW YORK STATE.
PA
XX
PI Barnwell JW, Longacre-Andre S, Mendis K, Nato F;
PI Roth C;
XX
DR WPI: 1997-425033/39.
DR N-PSDB; AAT94549.
XX
XX Recombinant protein containing the merozoite surface protein-1 p19
PT fragment - useful in anti-malarial vaccines, diagnosis and protein
PT purification
XX
XX Disclosure; Fig 1C; 85pp; French.
XX
CC This is the amino acid sequence of a chimeric protein comprising amino
CC acids 1-32 of the Plasmodium vivax merozoite surface protein 1 (MSp1)
CC linked to amino acids Asn1613-Ser1705 which correspond to the 19 kD
CC C-terminal fragment (p19) of MSP-1 from Plasmodium falciparum.
CC p19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium
CC species.
CC The recombinant protein can be used for the production of anti-malarial
CC vaccines, where the p19 fragment provides a high level of protective
CC immunity since it includes epitopes not presented in the p42 fragment.
XX
SQ Sequence 127 AA;

Query Match 100.0%; Score 105; DB 18; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
Db |||||
49 nsqcfrrhldereckcll 66

RESULT 10
ID AAY70278 standard; Protein; 350 AA.
XX
AC AAY70278;
XX
DT 06-JUN-2000 (first entry)
XX
DE Recombinant vaccine CDC/NIIMALVAC-1.
XX
KW Recombinant protein; CDC/NIIMALVAC-1; multivalent; malaria; vaccine;
KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
KW Pf27; antiparasitic; prevention; anti-CDC/NIIMALVAC-1 antibody;
KW honey bee.
XX
OS Chimeric - Apis sp.
OS Chimeric - Clostridium tetani.
OS Chimeric - Plasmodium falciparum.
XX
FH Key Location/Qualifiers
FT Peptide 1..22
FT /label= Melittin_signal_peptide
FT /note= "Derived from Honey bee"
FT 23..350
FT Protein /label= Mature_CDC/NIIMALVAC-1
FT /note= "Recombinant multivalent malarial vaccine"

```

Mon Feb 4 15:23:34 2002

XX PN WO200011179-A1.
XX PD 02-MAR-2000.
XX PF 19-AUG-1999; 99WO-US18869.
XX PR 21-AUG-1998; 98US-0097703.
XX PS (NIM-) NAT INST IMMUNOLOGY.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Lal AA, Shi YP, Hasnain SE;
XX DR WPI: 2000-237654/20.
XX DR N-PSDB; AAX51336.
XX PT Novel recombinant protein as vaccine for treating malarial infection
XX PT comprises antigenic peptides obtained from different stages of
XX PT plasmodium falciparum life cycle
XX PS Claim 3; Page 43-44; 52pp; English.
XX CC The present sequence is that of recombinant protein CDC/N1MALVAC-1,
XX CC which is a multivalent, multistage malarial vaccine. The recombinant
XX CC protein comprises melittin signal peptide, (His)6 tag, T-cell epitope
XX CC from tetanus toxoid and 21 antigenic epitopes from circumsporozoite
XX CC protein (CSP), sporozoite surface protein-1 (MSP-1), MSP-2, apical
XX CC antigen-1 (LSA-1), merozoite surface protein-2 (SSP-2), liver stage
XX CC membrane associated protein-1 (RAP-1) and gamete specific antigen, PfG27.
XX CC These epitopes were obtained at different stages of the life cycle of
XX CC plasmodium falciparum. CDC/N1MALVAC-1 vaccine has antiparasitic
XX CC activity and can be used for treatment and prevention of malarial
XX CC infections. Anti-CDC/N1MALVAC-1 antibodies can be used for detecting
XX CC P. falciparum in biological samples.
XX SQ Sequence 350 AA;

Query Match 100.0%; Score 105; DB 21; Length 350;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18
DB 82 nsgcfrhldereckcll 99
|||||

RESULT 11
AAY09372
ID AAY09372 standard; Protein; 355 AA.

XX AC AAY09372;

XX DT 31-AUG-1999 (first entry)

XX DE Merozoite surface protein MSP-1-42.

XX KW MSP-1; merozoite surface protein; malaria; vaccine;
XX KW protein engineering; protein expression; codon usage;
XX KW transgenic animal.

XX OS Plasmodium falciparum.

XX PN WO9920774-A2.

XX PD 29-APR-1999.

XX PF 20-OCT-1998; 98WO-US22226.

XX PR 15-MAY-1998; 98US-0085649.

XX PR 20-OCT-1997; 97US-0062592.

XX PA (GENZ) GENZYME TRANSGENICS CORP.
XX PI Chen LH, Meade H;
XX DR WPI: 1999-288313/24.
XX DR P-PSDB; AAX56008.
XX PT Modified malarial protein for use in anti-malarial vaccines
XX PS Example; Fig 1; 35pp; English.

XX CC The present sequence represents a 42 kDa C-terminal portion of
XX CC malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an
XX CC important target for the development of a vaccine against
XX CC plasmodium falciparum. The nucleic acid sequence encoding MSP-1-42
XX CC has been modified (see AAX56008) compared to the native sequence to
XX CC lower the AT content (from 76 to 49.7%) and to eliminate 10 mRNA
XX CC instability motifs while maintaining the same protein amino acid
XX CC sequence. These alterations allow MSP-1-42 to be expressed in
XX CC mammalian cell culture and in transgenic mice. Native MSP-1-12
XX CC is known to be difficult to express in cell culture systems,
XX CC mammalian cell culture systems or in transgenic animals. The
XX CC invention allows expression of MSP-1 protein in the milk of
XX CC transgenic animals, and also provides a DNA vaccine comprising a
XX CC vector containing the altered MSP-1-42 sequence.

XX SQ Sequence 355 AA;

Query Match 100.0%; Score 105; DB 20; Length 355;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18
DB 277 nsgcfrhldereckcll 294
|||||

RESULT 12
AAY05832
ID AAY05832 standard; Protein; 355 AA.

XX AC AAY05832;

XX DT 02-AUG-1999 (first entry)

XX DE Merozoite surface protein MSP-1-42.

XX KW MSP-1; merozoite surface protein; malaria; vaccine;
XX KW protein engineering; protein expression; codon usage;
XX KW transgenic animal.

XX OS Plasmodium falciparum.

XX PN WO9920766-A2.

XX PD 29-APR-1999.

XX PF 20-OCT-1998; 98WO-US22225.

XX PR 15-MAY-1998; 98US-0085649.

XX PR 20-OCT-1997; 97US-0062592.

XX PA (GENZ) GENZYME TRANSGENICS CORP.

XX PI Chen LH, Meade H;

XX DR WPI: 1999-302742/25.

XX DR N-PSDB; AAX25586.

XX PR New modified recombinant nucleic acid sequences useful for producing

PT malarial DNA vaccine
 PS Disclosure; Fig 1; 43pp; English.
 XX The present sequence represents a 42 kDa C-terminal portion of
 CC malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an
 CC important target for the development of a vaccine against
 CC Plasmodium falciparum. The nucleic acid sequence encoding MSP-1-42
 CC has been modified (see AAX25586) compared to the native sequence (see
 CC AAX25587) such that 306 nucleotide positions have been replaced to
 CC lower the AT content (from 76 to 49.7%) and to eliminate 10 mRNA
 CC instability motifs while maintaining the same protein amino acid
 CC sequence. These alterations allow MSP-1-42 to be expressed in
 CC mammalian cell culture and in transgenic mice. The invention
 CC provides modified recombinant nucleic acid sequences and methods for
 CC increasing the mRNA levels and protein expression of proteins that
 CC are difficult to express in cell culture systems, mammalian cell
 CC culture systems or in transgenic animals. The preferred difficult
 CC protein candidates for expression are those derived from lower
 CC organisms such as parasites, bacteria and viruses that have DNA
 CC coding sequences of high AT content or which have mRNA instability
 CC motifs or rare codons relative to the recombinant expression system
 CC to be used. The invention allows expression of MSP-1 protein in
 CC the milk of transgenic animals, and also provides a DNA vaccine
 CC comprising a vector containing the altered MSP-1-42 sequence.
 XX Sequence 355 AA;
 SQ

Query Match 100.0%; Score 105; DB 20; Length 355;
 Best Local Similarity 100.0%; Pred. No. 4, 4e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSGCFRHLDERECKLL 18
 |||||
 Db 277 nsgcfrhldececkll 294

RESULT 13
 AAY09373
 ID AAY09373 standard; Protein; 361 AA.
 AC AAY09373;
 XX 31-AUG-1999 (first entry)
 DT Merozoite surface protein MSP-1-42.
 DE MSP-1; merozoite surface protein; malaria; vaccine;
 KW protein engineering; protein expression; codon usage;
 KW transgenic animal.
 XX Plasmodium falciparum.
 OS WO9920774-A2.
 PN 29-APR-1999.
 PD 20-OCT-1998; 98WO-US222226.
 PF 15-MAY-1998; 98US-0085649.
 PR 20-OCT-1997; 97US-0062592.
 XX (GENZ) GENZYME TRANSGENICS CORP.
 PA Chen LH, Meade H;
 PI WPI: 1999-288313/24.
 XX P-PSDB; AAX56009.
 DR Modified malarial protein for use in anti-malarial vaccines
 PT Example; Fig 2; 35pp; English.
 XX

XX This present sequence comprises a 42 kDa C-terminal portion of
 CC malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an
 CC important target for the development of a vaccine against
 CC Plasmodium falciparum. The C-terminal end of the sequence is
 CC modified to include a 6xHis tag. A nucleic acid (see AAX56008)
 CC encoding MSP-1-42 has been modified according to a method
 CC of the invention in order to improve expression in mammalian cells
 CC and in transgenic animals by reducing the AT content and removing
 CC mRNA instability motifs. The invention allows expression of
 CC MSP-1-42 in the milk of transgenic animals, and also provides a DNA
 CC vaccine comprising a vector containing the altered MSP-1-42 nucleic
 CC acid.
 XX Sequence 361 AA;
 SQ

Query Match 100.0%; Score 105; DB 20; Length 361;
 Best Local Similarity 100.0%; Pred. No. 4, 5e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSGCFRHLDERECKLL 18
 |||||
 Db 277 nsgcfrhldececkll 294

RESULT 14
 AAY05833
 ID AAY05833 standard; Protein; 361 AA.
 AC AAY05833;
 XX 02-AUG-1999 (first entry)
 DT Merozoite surface protein MSP-1-42.
 DE MSP-1; merozoite surface protein; malaria; vaccine;
 KW protein engineering; protein expression; codon usage;
 KW transgenic animal.
 XX Plasmodium falciparum.
 OS WO9920766-A2.
 PN 29-APR-1999.
 PD 20-OCT-1998; 98WO-US222225.
 PF 15-MAY-1998; 98US-0085649.
 PR 20-OCT-1997; 97US-0062592.
 XX (GENZ) GENZYME TRANSGENICS CORP.
 PA Chen LH, Meade H;
 PI WPI: 1999-302742/25.
 XX N-PSDB; AAX25587.
 DR New modified recombinant nucleic acid sequences useful for producing
 PT malarial DNA vaccine
 XX Disclosure; Fig 2; 43pp; English.
 XX This present sequence comprises a 42 kDa C-terminal portion of
 CC malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an
 CC important target for the development of a vaccine against
 CC Plasmodium falciparum. The C-terminal end of the sequence is
 CC modified to include a 6xHis tag. Nucleic acids (see AAX25586 and
 CC AAX25593) encoding MSP-1-42 have been modified according to a method
 CC of the invention in order to improve expression in mammalian cells
 CC and in transgenic animals. The invention provides modified
 CC recombinant nucleic acid sequences and methods for increasing the
 CC mRNA levels and protein expression of proteins that are difficult

CC to express in cell culture systems, especially mammalian cell
 CC culture systems or in transgenic animals. The preferred difficult
 CC protein candidates for expression are those derived from lower
 CC organisms such as parasites, bacteria and viruses that have DNA
 CC coding sequences of high AT content or which have mRNA instability
 CC motifs or rare codons relative to the recombinant expression system
 CC to be used. The invention allows expression of MSP-1 in the milk
 CC of transgenic animals, and also provides a DNA vaccine comprising a
 CC vector containing the altered MSP-1-42 nucleic acid.

XX Sequence 361 AA;

Query Match 100.0%; Score 105; DB 20; Length 361;

Best Local Similarity 100.0%; Pred. No. 4.5e-07;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
 |||||
 DB 277 nsgcfhrldereckcll 294

RESULT 15

AAB83926
 ID AAB83926 standard; Protein; 375 AA.

XX
 AC AAB83926;

XX
 DT 23-JUL-2001 (first entry)

XX
 DE A major merozoite surface protein-1 fragment of 42kDa.

XX
 KW Major merozoite surface protein-1; MSP1-42; melittin signal peptide;
 malaria vaccine.

XX
 OS Plasmodium falciparum.

XX
 PN WO200134188-A1.

XX
 PD 17-MAY-2001.

XX
 PF 09-NOV-2000; 2000WO-US31064.

XX
 PR 12-NOV-1999; 99US-0165178.

XX
 PR 01-DEC-1999; 99US-0168327.

XX
 PR 22-AUG-2000; 2000US-0226861.

XX
 PA (UYHA-) UNIV HAWAII.

XX
 PA (UYCH-) UNIV CHINESE HONG KONG.

XX
 PA (QUEE-) QUEEN EMMA FOUND.

XX
 PI Hui GSN, Lap-Yin P, Ho WKK;

XX
 DR WPI; 2001-335879/35.

XX
 DR N-PSDB; AAF89840.

XX
 PT Producing malaria vaccine, useful for treatment or prevention of all
 PT forms of malaria in humans, by expressing immunogenic merozoite protein
 PT fragment in a baculovirus system -

XX
 PS Example 3; Page 87-88; 95pp; English.

XX
 CC The present sequence represents a major merozoite surface protein-1
 CC C-terminal fragment of 42kDa (MSP1-42). This fragment is linked to a
 CC melittin signal peptide, and then expressed in a in a
 CC silkworm/baculovirus system. The protein is used to prepare a
 CC malaria vaccine, which is used to treat or prevent malaria, caused by
 CC any of the four species of Plasmodium that infect humans.

XX
 SQ Sequence 375 AA;

Query Match 100.0%; Score 105; DB 22; Length 375;

Best Local Similarity 100.0%; Pred. No. 4.7e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
 |||||
 DB 296 nsgcfhrldereckcll 313

Search completed: January 29, 2002, 10:21:45
 Job time: 421 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:24:04 ; Search time 133.18 Seconds
(without alignments)
3.041 Million cell updates/sec

Title: US-09-763-397A-11

Perfect score: 105

Sequence: 1 NSGCFRHLDERECKLL 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCFUS-COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	100.0	48	1 US-08-290-919-1	Sequence 1, Appl
2	105	100.0	48	1 US-08-290-919-2	Sequence 2, Appl
3	105	100.0	48	1 US-08-290-919-12	Sequence 12, Appl
4	53	50.5	106	1 US-08-290-919-11	Sequence 11, Appl
5	45	42.9	2485	5 PCT-US94-00198-1	Sequence 1, Appl
6	45	42.9	2485	5 PCT-US94-00198-2	Sequence 2, Appl
7	45	42.9	2818	1 US-08-510-284-1	Sequence 1, Appl
8	45	42.9	2818	1 US-08-411-389-2	Sequence 2, Appl
9	45	42.9	2818	2 US-08-449-933-2	Sequence 2, Appl
10	45	42.9	2818	4 US-07-966-049A-2	Sequence 2, Appl
11	45	42.9	2818	4 US-09-542-331-2	Sequence 2, Appl
12	44	41.9	27	1 US-08-619-936-1	Sequence 1, Appl
13	44	41.9	567	2 US-08-841-483-2	Sequence 2, Appl
14	44	41.9	567	4 US-09-382-911-2	Sequence 2, Appl
15	43	41.0	236	1 US-08-015-985-6	Sequence 6, Appl
16	43	41.0	254	2 US-08-685-992-9	Sequence 9, Appl
17	43	41.0	254	2 US-09-144-925-9	Sequence 9, Appl
18	43	41.0	703	1 US-08-015-985-3	Sequence 3, Appl
19	43	41.0	802	1 US-08-015-985-1	Sequence 1, Appl
20	42	40.0	82	1 US-07-965-674-1	Sequence 1, Appl
21	42	40.0	82	5 PCT-US93-09523-1	Sequence 1, Appl
22	41	39.0	753	4 US-08-942-686-2	Sequence 2, Appl
23	40.5	38.6	150	4 US-08-334-179A-6	Sequence 6, Appl
24	40.5	38.6	150	4 US-08-334-179A-10	Sequence 10, Appl
25	40.5	38.6	578	4 US-08-981-392-13	Sequence 13, Appl
26	40.5	38.6	582	4 US-08-334-179A-4	Sequence 4, Appl
27	40.5	38.6	1038	4 US-08-334-179A-2	Sequence 2, Appl

28	40.5	38.6	1038	4 US-08-334-179A-8	Sequence 8, Appl
29	40	38.1	76	2 US-08-480-478-55	Sequence 55, Appl
30	40	38.1	76	2 US-08-326-110A-55	Sequence 55, Appl
31	40	38.1	77	2 US-08-465-380-55	Sequence 55, Appl
32	40	38.1	77	2 US-08-486-397-55	Sequence 55, Appl
33	40	38.1	77	2 US-08-486-399-55	Sequence 55, Appl
34	40	38.1	77	2 US-08-461-965-55	Sequence 55, Appl
35	40	38.1	77	2 US-08-634-641-55	Sequence 55, Appl
36	40	38.1	77	3 US-09-249-471-55	Sequence 55, Appl
37	40	38.1	77	3 US-09-249-472-55	Sequence 55, Appl
38	40	38.1	77	3 US-09-249-451-55	Sequence 55, Appl
39	40	38.1	77	3 US-08-809-455-55	Sequence 55, Appl
40	40	38.1	77	3 US-08-249-461-55	Sequence 55, Appl
41	40	38.1	77	3 US-09-249-448-55	Sequence 55, Appl
42	39.5	37.6	886	3 US-09-110-116-3	Sequence 3, Appl
43	39	37.1	49	2 US-08-841-483-32	Sequence 32, Appl
44	39	37.1	49	4 US-09-382-911-32	Sequence 32, Appl
45	38.5	36.7	721	3 US-08-872-855-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-290-919-1
; Sequence 1, Application US/08290919
; Patent No. 5720959
; GENERAL INFORMATION:
; APPLICANT: HOLDER, ANTHONY A.
; APPLICANT: BLACKMAN, MICHAEL J.
; APPLICANT: CHAPPEL, JONATHAN A.
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
; TITLE OF INVENTION: VACCINE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290.919
; FILING DATE: 04-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9203821.5
; FILING DATE: 22-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/00367
; FILING DATE: 22-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 212242/HCM/NJL/6BC8/
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site

us-09-763-397a-11.ra1

Mon Feb 4 15:23:34 2002

LOCATION: 1
OTHER INFORMATION: /label= X
OTHER INFORMATION: /note= "X" = M and N, or N"
US-08-290-919-1

Query Match 100.0%; Score 105; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18
Db 15 NSGCFRHLDERECKLL 32

RESULT 2
US-08-290-919-2
Sequence 2, Application US/08290919
Patent No. 5720959
GENERAL INFORMATION:
APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
TITLE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,919
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9203821.5
FILING DATE: 22-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00367
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 212242/HCM/MJL/6BC8/
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label= X
OTHER INFORMATION: /note= "X" = M and N, or N"
US-08-290-919-2

Query Match 100.0%; Score 105; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NSGCFRHLDERECKLL 18
Db 15 NSGCFRHLDERECKLL 32

RESULT 3
US-08-290-919-12
Sequence 12, Application US/08290919
Patent No. 5720959
GENERAL INFORMATION:
APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
TITLE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,919
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9203821.5
FILING DATE: 22-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00367
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 212242/HCM/MJL/6BC8/
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-290-919-12

Query Match 100.0%; Score 105; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18
Db 15 NSGCFRHLDERECKLL 32

RESULT 4
US-08-290-919-11
Sequence 11, Application US/08290919
Patent No. 5720959
GENERAL INFORMATION:
APPLICANT: HOLDER, ANTHONY A.

```

: STREET: 1 Girard Farms
:
: CITY: Madison
:
: STATE: New Jersey
:
: COUNTRY: USA
: ZIP: 94304-1104
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: Macintosh
:
: OPERATING SYSTEM: 6.0.8
:
: SOFTWARE: Microsoft Word 5.1a
:

```

```

1  RESULT      6
2  PCT-US94-001198-2
3  1 Sequence 2, Application PC/TUS94001198
4  GENERAL INFORMATION:
5  APPLICANT: Schering Corp.
6  TITLE OF INVENTION: HAS Associated GAP Proteins
7  NUMBER OF SEQUENCES: 6
8  CORRESPONDENCE ADDRESS:
9  ADDRESSEE: Schering Corp.
10 STREET: 1 Girald Farms
11 CITY: Madison
12 STATE: New Jersey
13 COUNTRY: USA
14 ZIP: 94304-1104
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Floppy disk
17 COMPUTER: Macintosh
18 OPERATING SYSTEM: 6.0.8
19 SOFTWARE: Microsoft Word 5.1a
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: PCT/US94/001198
22 FILING DATE:
23 CLASSIFICATION:
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: US 08/004,824
26 FILING DATE: 15-JAN-1993
27 ATTORNEY/AGENT INFORMATION:
28 NAME: Lunn, Paul G.
29 REGISTRATION NUMBER: 32,743
30 REFERENCE/DOCKET NUMBER: DX0352 PCT
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: (201)822-7255
33 TELEFAX: (201)822-7039
34 INFORMATION FOR SEQ ID NO: 2:
35 SEQUENCE CHARACTERISTICS:

```

; LENGTH: 2485 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
PCT-US94-00198-2

Query Match 42.9%; Score 45; DB 5; Length 2485;
Best Local Similarity 53.8%; Pred. No. 2.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 4 CFRHLDERECKC 16
Db 375 CFRHCEADIRC 387
||||| : :|

RESULT 7

US-08-510-284-1
; Sequence 1, Application US/08510284
; Patent No. 5580955

; GENERAL INFORMATION:
; APPLICANT: Nur-E-Kamal, M. S. A.; Maruta, Hiroshi

; TITLE OF INVENTION: FRAGMENTS OF NEUROFIBROMIN (NFI) AND METHOD
; TITLE OF INVENTION: TO REVERSE ACTIVATED RAS INDUCED MALIGNANT TRANSFORMATION IN
; TITLE OF INVENTION: MAMMALIAN CELLS

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe & Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/510,284

; FILING DATE:

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/071,575

; FILING DATE: 1-JUNE-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Hanson, No. 5580955man D.

; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: LUD 5319

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2818 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; PUBLICATION INFORMATION:

; AUTHORS: Marchuk, Douglas A.; Saulino, Ann M.;

; AUTHORS: TavaKKol, Roxanne; Swaroop, Manju;

; AUTHORS: Wallace, Margaret R.; Andersen, Lone B.;

; AUTHORS: Mitchell, Anna L.; Gutmann, David H.;

; AUTHORS: Boguski, Mark; Collins, Francis S.

; TITLE: cDNA Cloning of the Type 1 Neurofibromatosis Gene;

; TITLE: Complete Sequence of the NFI Gene Product

; Patent No. 5580955

; JOURNAL: Genomics

; VOLUME: 11

; PAGES: 931-940

; DATE: 1991
US-08-510-284-1

Query Match 42.9%; Score 45; DB 1; Length 2818;
Best Local Similarity 53.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 4 CFRHLDERECKC 16
Db 709 CFRHCEADIRC 721
||||| : :|

RESULT 8

US-08-411-389-2
; Sequence 2, Application US/08411389
; Patent No. 5605799

; GENERAL INFORMATION:
; APPLICANT: White, Raymond L.

; APPLICANT: Cawthon, Richard M.

; APPLICANT: Li, Ying

; TITLE OF INVENTION: SOMATIC MUTATIONS IN THE

; TITLE OF INVENTION: NEUROFIBROMATOSIS TYPE 1 GENE IN HUMAN TUMORS

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Venable, Baetjer, Howard & Civiletti

; STREET: 1201 New York Avenue NW, Suite 1000

; CITY: Washington

; STATE: DC

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/411,389

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/047,088

; FILING DATE: 16-APR-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Ihnen, Jeffrey L.

; REGISTRATION NUMBER: 28,957

; REFERENCE/DOCKET NUMBER: 19780-107116

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-962-8300

; TELEFAX: 202-962-8310

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2818 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-411-389-2

Query Match 42.9%; Score 45; DB 1; Length 2818;
Best Local Similarity 53.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 4 CFRHLDERECKC 16
Db 709 CFRHCEADIRC 721
||||| : :|

RESULT 9

US-08-449-933-2
; Sequence 2, Application US/08449933
; Patent No. 5859195

; GENERAL INFORMATION:
; APPLICANT: Collins, Francis S.

APPLICANT: Wallace, Margaret R.
 APPLICANT: Marchuk, Douglas A.
 APPLICANT: Anderson, Lone B.
 APPLICANT: Guttman, David H.
 TITLE OF INVENTION: Neurofibromatosis Gene
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 755 Page Mill Road
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304-1018
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/449,933
 FILING DATE: 25-MAY-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Kanski, Antoinette F.
 REGISTRATION NUMBER: 34,202
 REFERENCE/DOCKET NUMBER: 20344-20553.10
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 813-5600
 TELEFAX: (415) 494-0792
 TELEX: 706141 MRSNFOERS SFO
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2818 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 POSITION IN GENOME:
 CHROMOSOME/SEGMENT: 17q11.2
 FEATURE:
 NAME/KEY: Cleavage-site
 LOCATION: group(583..586, 815..818, 2573..2576, 2810..2813)
 OTHER INFORMATION: /note= "Potential CAMP-dependent
 OTHER INFORMATION: protein kinase recognition sites"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 2549..2556
 OTHER INFORMATION: /note= "Potential tyrosine
 OTHER INFORMATION: phosphorylation site"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: group(1264..1276, 1358, 1377, 1389, 1390, 1391,
 LOCATION: 1395, 1396, 1400, 1423, 1426, 1429, 1430)
 OTHER INFORMATION: /note= "Invariant residues within
 OTHER INFORMATION: most statistically significant regions of similarity among th
 OTHER INFORMATION: GAP family of proteins"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: group(1264..1290, 1345..1407, 1415..1430)
 OTHER INFORMATION: /note= "Most statistically
 OTHER INFORMATION: significant regions of similarity among the GAP family of
 OTHER INFORMATION: proteins"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 496
 OTHER INFORMATION: /note= "At variance with previously
 OTHER INFORMATION: published sequence which shows an ATG methionine codon rather
 OTHER INFORMATION: than an ATA isoleucine codon"
 FEATURE:

NAME/KEY: Modified-site
 LOCATION: 1183
 OTHER INFORMATION: /note= "At variance with previously
 OTHER INFORMATION: published sequence. Shows an CTG leucine codon rather than
 OTHER INFORMATION: previously published CTC"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 1555
 OTHER INFORMATION: /note= "At variance with previously
 OTHER INFORMATION: published sequence. Lacks an extra CAT histidine condon af
 OTHER INFORMATION: this residue"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: (2771~2772)
 OTHER INFORMATION: /note= "Position of an 18 amino
 OTHER INFORMATION: acid insertion(SEQ ID NO:10) representing an alternatively
 OTHER INFORMATION: spliced product"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: (1370~1371)
 OTHER INFORMATION: /note= "Position of a 21 amino acid
 OTHER INFORMATION: insertion representing an alternatively spliced product"
 FEATURE:
 NAME/KEY: Domain
 LOCATION: 1125..1537
 OTHER INFORMATION: /note= "Nf1 catalytic domain"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 2746..2818
 OTHER INFORMATION: /note= "Corresponding amino acids
 OTHER INFORMATION: for the PstI-HindIII fragment designated pMAL.B3A"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 65..371
 OTHER INFORMATION: /note= "Corresponding amino acids
 OTHER INFORMATION: for the HpaI-PstI fragment designated pMAL.HF3A.P"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 65..1240
 OTHER INFORMATION: /note= "Corresponding amino acids
 OTHER INFORMATION: for the HpaI-XhoI fragment designated pMAL.HF3A.X"
 PUBLICATION INFORMATION:
 AUTHORS: Wallace, M.R. et al.
 TITLE: Type 1 Neurofibromatosis Gene: Correction
 JOURNAL: Science
 VOLUME: 250
 ISSUE: 12/21/90
 PAGES: 1749-
 DATE: 12/21-1990
 RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 2818
 PUBLICATION INFORMATION:
 AUTHORS: Wallace, M.R. et al.
 TITLE: Type 1 Neurofibromatosis Gene: Identification
 TITLE: of a Large Transcript in Three Nf1 Patients
 JOURNAL: Science
 VOLUME: 249
 ISSUE: 07/13/90
 PAGES: 181-186
 DATE: 07/13-1990
 RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 2818
 US-08-449-933-2

Query Match 42.9%; Score 45; DB 2; Length 2818;
 Best Local Similarity 53.8%; Pred. No. 2.4e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 4 CFRHLDERECKC 16

Db 709 CFRHLCEADIRC 721

RESULT 10

```
US-07-966-049A-2
; Sequence 2, Application US/07966049A
; Patent No. 6238861
; GENERAL INFORMATION:
; APPLICANT: Collins, Francis S.
; APPLICANT: Wallace, Margaret R.
; APPLICANT: Marchuk, Douglas A.
; APPLICANT: Anderson, Lone B.
; APPLICANT: Guttman, David H.
; TITLE OF INVENTION: Neurofibromatosis Gene
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/07/966,049A
; FILING DATE: 02-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Konski, Antoinette F.
; REGISTRATION NUMBER: 34,202
; REFERENCE/DOCKET NUMBER: 20344-20553.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141 MRSNFOERS SFO
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2818 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 17q11.2
; FEATURE:
; NAME/KEY: Cleavage-site
; LOCATION: group(1264, 1276, 1358, 1377, 1389, 1390, 1391,
; OTHER INFORMATION: /note= "Potential tyrosine
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2549..2556
; OTHER INFORMATION: /note= "Potential tyrosine
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1395, 1396, 1400, 1423, 1426, 1429, 1430)
; OTHER INFORMATION: /note= "Invariant residues within
; OTHER INFORMATION: most statistically significant regions of similarity among th
; OTHER INFORMATION: GAP family of proteins"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: group(1264..1290, 1345..1407, 1415..1430)
; OTHER INFORMATION: /note= "Most statistically
; OTHER INFORMATION: significant regions of similarity among the GAP family of
; OTHER INFORMATION: proteins"
; FEATURE:
; NAME/KEY: Modified-site

; LOCATION: 496
; OTHER INFORMATION: /note= "At variance with previously
; OTHER INFORMATION: published sequence which shows an ATG methionine codon rat
; OTHER INFORMATION: than an ATA isoleucine codon"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1183
; OTHER INFORMATION: /note= "At variance with previously
; OTHER INFORMATION: published sequence. Shows an CTG leucine codon rather than
; OTHER INFORMATION: previously published CTC"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1555
; OTHER INFORMATION: /note= "At variance with previously
; OTHER INFORMATION: published sequence. Lacks an extra CAT histidine condon af
; OTHER INFORMATION: this residue"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: (2771^2772)
; OTHER INFORMATION: /note= "Position of an 18 amino
; OTHER INFORMATION: acid insertion(SEQ ID NO:10) representing an alternatively
; OTHER INFORMATION: spliced product"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: (1370^1371)
; OTHER INFORMATION: /note= "Position of a 21 amino acid
; OTHER INFORMATION: insertion representing an alternatively spliced product"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 1125..1537
; OTHER INFORMATION: /note= "NFL catalytic domain"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2746..2818
; OTHER INFORMATION: /note= "Corresponding amino acids
; OTHER INFORMATION: for the PstI-HindIII fragment designated pMAL.B3A"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 65..371
; OTHER INFORMATION: /note= "Corresponding amino acids
; OTHER INFORMATION: for the HpaI-PstI fragment designated pMAL.HF3A.P"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 65..1240
; OTHER INFORMATION: /note= "Corresponding amino acids
; OTHER INFORMATION: for the HpaI-XhoI fragment designated pMAL.HF3A.X"
; PUBLICATION INFORMATION:
; AUTHORS: Wallace, M.R. et al.
; TITLE: Type 1 Neurofibromatosis Gene: Correction
; JOURNAL: Science
; VOLUME: 250
; ISSUE: 12/21/90
; PAGES: 1749-
; DATE: 12/21-1990
; RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 2818
; PUBLICATION INFORMATION:
; AUTHORS: Wallace, M.R. et al.
; TITLE: Type 1 Neurofibromatosis Gene: Identification
; JOURNAL: Science
; VOLUME: 249
; ISSUE: 07/13/90
; PAGES: 181-186
; DATE: 07/13-1990
; RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 2818
; US-07-966-049A-2

Query Match 42.9%; Score 45; DB 4; Length 2818;
Best Local Similarity 53.8%; Pred. No. 2.4e-02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 4 CFRHLDERECKC 16
```



```
Db 709 CFRHLCCEADIRC 721
||||| | : |
RESULT 11
US-09-542-331-2
; Sequence 2, Application US/09542331
; Patent No. 6261761
; GENERAL INFORMATION:
; APPLICANT: Zhong, Yi
; TITLE OF INVENTION: NF1 Protein and Its Role in Activation
; FILE REFERENCE: 1314.1047003
; CURRENT APPLICATION NUMBER: US/09/542,331
; EARLIER FILING DATE: 2000-04-04
; EARLIER APPLICATION NUMBER: US 09/046,745
; EARLIER FILING DATE: 1998-03-24
; EARLIER APPLICATION NUMBER: US 60/041,469
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2818
; TYPE: PRT
; ORGANISM: Human
US-09-542-331-2

Query Match 42.9%; Score 45; DB 4; Length 2818;
Best Local Similarity 53.8%; Pred. No. 2 4e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 CFRHLDRECEKC 16
||||| | : |
Db 709 CFRHLCCEADIRC 721

RESULT 12
US-08-619-936-1
; Sequence 1, Application US/08619936
; Patent No. 5672682
; GENERAL INFORMATION:
; APPLICANT: Terlau, Heinrich
; APPLICANT: Shon, Ki-Joon
; APPLICANT: Grilley, Michelle
; APPLICANT: Olivera, Baldomero M.
; TITLE OF INVENTION: Conotoxin Peptide PVIIA
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
; STREET: 1201 New York Avenue, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,936
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Saxe, Stephen A.
; REGISTRATION NUMBER: 38,609
; REFERENCE/DOCKET NUMBER: 24260-107674-04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4848
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus purpurascens
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4
; OTHER INFORMATION: /product= "4Hyp"
; OTHER INFORMATION: /note= "Amino acid 4 may be 4-trans-hydroxyproline."
; NAME/KEY: Disulfide-bond
; LOCATION: 1..16
; FEATURE:
; NAME/KEY: Disulfide-bond
; LOCATION: 8..20
; FEATURE:
; NAME/KEY: Disulfide-bond
; LOCATION: 15..26
; US-08-619-936-1

Query Match 41.9%; Score 44; DB 1; Length 27;
Best Local Similarity 50.0%; Pred. No. 3,3;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECK 16
| | | | | |
Db 5 NQKCFQHLDDCCSRKC 20

RESULT 13
US-08-841-483-2
; Sequence 2, Application US/08841483B
; Patent No. 5976875
; GENERAL INFORMATION:
; APPLICANT: Prescott, Steven M.
; APPLICANT: Bunting, Michaeline
; APPLICANT: Tang, Wen
; APPLICANT: Topham, Matthew
; TITLE OF INVENTION: Diacylglycerol Kinase Isoforms Epsilon and Zeta and
; FILE REFERENCE: 2037.21a
; CURRENT APPLICATION NUMBER: US/08/841,483B
; EARLIER FILING DATE: 1997-04-22
; EARLIER APPLICATION NUMBER: 60/016,210
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-841-483-2

Query Match 41.9%; Score 44; DB 2; Length 567;
Best Local Similarity 46.7%; Pred. No. 67;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECK 15
| | | | | |
Db 95 DEGCLRKADKRFQCK 109

RESULT 14
US-09-382-911-2
; Sequence 2, Application US/09382911
; Patent No. 6221658
```

```

? GENERAL INFORMATION:
? APPLICANT: Prescott, Steven M.
? APPLICANT: Bunting, Michaelaine
? APPLICANT: Tang, Wen
? APPLICANT: Topham, Matthew
? TITLE OF INVENTION: Diacylglycerol Kinase Isoforms Epsilon and Zeta and
? TITLE OF INVENTION: Methods of Use Thereof
? FILE REFERENCE: 2037.2.1a
? CURRENT APPLICATION NUMBER: US/09/382.911
? PRIOR FILING DATE: 1999-08-25
? CURRENT APPLICATION NUMBER: 08/841.483
? PRIOR FILING DATE: 1997-04-22
? PRIOR APPLICATION NUMBER: 60/016.210
? PRIOR FILING DATE: 1996-04-22
? NUMBER OF SEQ ID NOS: 33
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 2
? LENGTH: 567
? TYPE: PRT
? ORGANISM: Homo sapiens
? OS-09-382-911-2

```

```

Query Match      41.9%; Score 44; DB 4; Length 567;
Best Local Similarity 46.7%; Pred. NO. 67;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECK 15
      : || | 1 : ||
Db 95 DEGCLRKADKRFCK 109

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:26:39 ; Search time 144.96 Seconds
(without alignments)
9.459 Million cell updates/sec

Title: US-09-763-397A-11

Perfect score: 105

Sequence: 1 NSGCFRHLDERECKLL 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	100.0	400	2 A45545	major merozoite su
2	105	100.0	651	2 SA7282	merozoite surface
3	105	100.0	1631	1 SAZQK1	major merozoite su
4	105	100.0	1639	2 S05603	major merozoite su
5	105	100.0	1701	2 A54498	major merozoite su
6	105	100.0	1701	2 A26868	major merozoite su
7	105	100.0	1726	1 SAZQGM	major merozoite su
8	105	100.0	1726	2 A45948	major merozoite su
9	100	95.2	1640	2 A24594	probable major sur
10	61	58.1	1785	2 A45546	major merozoite su
11	60	57.1	1726	2 A39401	merozoite surface
12	60	57.1	1751	2 A45604	major blood-stage
13	53	50.5	680	2 A28121	major merozoite su
14	53	50.5	1772	2 A45532	major merozoite su
15	46	43.8	298	2 F82532	isopentenyl monoph
16	46	43.8	560	2 C41539	kinesin light chai
17	46	43.8	569	2 I53013	kinesin light chai
18	45.5	43.3	479	2 A61368	keratin type II, h
19	45.5	43.3	495	2 I57463	keratin type II -
20	45	42.9	449	2 S07714	T64 protein precu
21	45	42.9	451	2 I50131	clusterin - quail
22	45	42.9	2818	2 S55282	neurofibromatosis
23	45	42.9	2820	2 JC5196	neurofibromin I -
24	45	42.9	2825	2 I54352	neurofibromin - mo
25	45	42.9	13055	2 T16580	hypothetical prote
26	44.5	42.4	116	1 S15197	hypA protein - Esc
27	44.5	42.4	116	2 F85921	pleiotropic effec
28	44.5	42.4	2109	1 I50421	aggreacan precursor
29	44	41.9	27	2 A58997	kappa-conotoxin pv

30 44 41.9 1444 2 B84809 hypothetical prote
31 43 41.0 796 1 JC1285 protein-tyrosine-p
32 43 41.0 802 1 A36065 protein-tyrosine-p
33 43 41.0 829 1 A47373 protein-tyrosine-p
34 42 40.0 775 2 J01429 hypothetical 8.1K
35 42 40.0 386 2 T27079 hypothetical prote
36 42 40.0 446 2 A42108 clusterin precursor
37 41.5 39.5 623 2 T47542 spot 3 protein and
38 41.5 39.5 1127 2 T25804 hypothetical prote
39 41 39.0 191 2 T19210 hypothetical prote
40 41 39.0 278 2 T16494 hypothetical prote
41 41 39.0 297 2 T36724 probable membrane
42 41 39.0 387 1 S26053 chormismate mutase
43 41 39.0 439 2 A35744 clusterin precursor
44 41 39.0 506 2 T50211 WD-repeat protein
45 41 39.0 776 2 S28258 androgen-regulated

RESULT 1
A45545
major merozoite surface antigen - malaria parasite (Plasmodium falciparum) (fragments
C:Species: Plasmodium falciparum
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jun-2000
C:Accession: A45545
R:Blackman, M.J.; Ling, I.T.; Nicholls, S.C.; Holder, A.A.
Mol. Biochem. Parasitol. 49, 29-33, 1991
A:Title: Proteolytic processing of the Plasmodium falciparum merozoite surface protei
A:Reference number: A45545; MUID:92131048
A:Accession: A45545
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-400 <BLA>
A>Note: sequence extracted from NCBI backbone (NCBIN:77612, NCBIP:77621)
C:Superfamily: major merozoite surface antigen
C:Keywords: glycoprotein; merozoite; surface antigen; tandem repeat

Query Match 100.0%; Score 105; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NSGCFRHLDERECKLL 18
Db 301 NSGCFRHLDERECKLL 318

RESULT 2
SA7282
merozoite surface antigen 1 - malaria parasite (Plasmodium falciparum) (strain RO-71)
C:Species: Plasmodium falciparum
A:Variety: strain RO-71
C:Date: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 09-Jun-2000
C:Accession: SA7282
R:Tolle, R.; Bujard, H.; Cooper, J.A.
submitted to the EMBL data library, July 1994
A:Description: Plasmodium falciparum: recombination within the C-terminal region of m
A:Reference number: SA7282
A:Accession: SA7282
A:Molecule type: DNA
A:Residues: 1-651 <TOI>
A:Cross-references: EMBL:Z35329; NID:9535257; PIDN:CAA84558.1; PID:g535258
A:Experimental source: strain RO-71
C:Superfamily: major merozoite surface antigen
C:Keywords: glycoprotein; merozoite; surface antigen

Query Match 100.0%; Score 105; DB 2; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.5e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ALIGNMENTS

R:Weber, J.L.; Sim, B.K.L.; Lyon, J.A.; Wolff, R.
Nucleic Acids Res. 16, 1206, 1988
A:Title: Merozoite surface protein sequence from the Camp strain of the human malaria parasite
A:Reference number: S06361; MUID:88143999
A:Accession: S06361
A:Molecule type: DNA
A:Residues: 1104-1726 <WEB2>
A:CROSS-references: EMBL:X03831
C:Comment: The merozoite stages of different strains have strain-specific surface antigens
C:Species: Plasmodium falciparum has three stages: sporozoite, merozoite, and gametocyte. The merozoite stage is the most abundant in the blood.
C:Superfamily: major merozoite surface antigen
C:Keywords: glycoprotein; merozoite; surface antigen; tandem repeat
F:1-19/domain: signal sequence #status predicted <SIG>
F:20-1726/product: major merozoite surface antigen #status predicted <MAT>
F:67-87/91-96/100-105/109-120/Region: 3-residue repeats (S-G-T)
F:757-765/Region: 3-residue repeats (T-E-E)
F:133,272,501,567,638,827,839,924,944,990,1016,1114,1221,1613,1658/Binding site: carboxy

Query Match 100.0%; Score 105; DB 1; Length 1726;
Best Local Similarity 100.0%; Pred. No. 8.4e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18
DB 1627 NSGCFRHLDERECKLL 1644

RESULT 8
A45948
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (sp)
C:Species: Plasmodium falciparum
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jun-2000
C:Accession: A45948
R:Chang, S.P.; Kramer, K.J.; Yamaga, K.M.; Kato, A.; Case, S.E.; Siddiqui, W.A.
Exp. Parasitol. 67, 1-11, 1988
A:Title: Plasmodium falciparum: gene structure and hydropathy profile of the major merozoite surface antigen
A:Reference number: A45948; MUID:89005525
A:Accession: A45948
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-1726 <CHA>
A:CROSS-references: GB:M37213
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen

Query Match 100.0%; Score 105; DB 2; Length 1726;
Best Local Similarity 100.0%; Pred. No. 8.4e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18
DB 1627 NSGCFRHLDERECKLL 1644

RESULT 9
A24594
probable major surface antigen (83K, 19K, 42K) precursor - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 29-Aug-1987 #sequence_revision 29-Aug-1987 #text_change 09-Jun-2000
C:Accession: A24594
R:Holder, A.A.; Lockyer, M.J.; Odink, K.G.; Sandhu, J.S.; Riveros-Moreno, V.; Nicholls, Nature 317, 270-273, 1985
A:Title: Primary structure of the precursor to the three major surface antigens of Plasmodium falciparum
A:Reference number: A24594; MUID:86014355
A:Accession: A24594
A:Molecule type: DNA
A:Residues: 1-1640 <HOL>
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen

Query Match 95.2%; Score 100; DB 2; Length 1640;
Best Local Similarity 94.4%; Pred. No. 4.5e-07;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18
DB 1541 NSGCFRHLDERECKLL 1558

RESULT 10
A45546
major merozoite surface antigen precursor - Plasmodium chabaudi chabaudi
C:Species: Plasmodium chabaudi chabaudi
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 26-Aug-1999
C:Accession: A45546
R:Deleersnijder, W.; Hendrix, D.; Bendahman, N.; Hanegreffe, J.; Brijs, L.; Hamers-Ca Mol. Biochem. Parasitol. 43, 231-244, 1990
A:Title: Molecular cloning and sequence analysis of the gene encoding the major merozoite surface antigen
A:Reference number: A45546; MUID:91218805
A:Accession: A45546
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1785
A:CROSS-references: GB:M34947; NID:gi160597; PID:gi160598
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen

Query Match 58.1%; Score 61; DB 2; Length 1785;
Best Local Similarity 61.1%; Pred. No. 0.35;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18
DB 1682 NAGCFRYDDGTTEWRCLL 1699

RESULT 11
A39401
merozoite surface antigen 1 precursor - Plasmodium vivax
C:Species: Plasmodium vivax
C:Date: 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 24-Nov-1999
C:Accession: A39401
R:Del Portillo, H.A.; Longacre, S.; Khouri, E.; David, P.H.
Proc. Natl. Acad. Sci. U.S.A. 88, 4030-4034, 1991
A:Title: Primary structure of the merozoite surface antigen 1 of Plasmodium vivax rev
A:Reference number: A39401; MUID:91219506
A:Accession: A39401
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1726
A:CROSS-references: GB:M60807
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen

Query Match 57.1%; Score 60; DB 2; Length 1726;
Best Local Similarity 55.6%; Pred. No. 0.48;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18
DB 1630 NAACYRYLDGMEWRCLL 1647

RESULT 12
A45604
major blood-stage surface antigen Pv200 - Plasmodium vivax
C:Species: Plasmodium vivax
C:Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999
C:Accession: A45604
R:Gibson, H.L.; Tucker, J.E.; Kaslow, D.C.; Kretzli, A.U.; Collins, W.E.; Kieffer, M.C Mol. Biochem. Parasitol. 50, 325-333, 1992

A:Title: Structure and expression of the gene for Pv200, a major blood-stage surface antigen
 A:Reference number: A45604; MUID:92158013
 A:Accession: A45604
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1751 <GIB>
 A:Cross-references: GB:M75674; NID:gl60608; PID:gl57336
 A:Note: sequence extracted from NCBI backbone (NCBIN:83591, NCBI:83592)
 C:Superfamily: major merozoite surface antigen
 C:Keywords: surface antigen

Query Match 57.1%; Score 60; DB 2; Length 1751;
 Best Local Similarity 55.6%; Pred. No. 0.49;
 Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
 I:||||| : |||

DB 1655 NAACYRYLDGTGEWRCLL 1672

RESULT 13

A28121
 A:Title: Cloning and analysis of the gene for a Plasmodium yoelii merozoite surface antigen
 C:Species: Plasmodium yoelii
 C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 26-Aug-1999
 C:Accession: A28121
 R:Burns Jr., J.M.; Daly, T.M.; Vaidya, A.B.; Long, C.A.
 Proc. Natl. Acad. Sci. U.S.A. 85, 602-606, 1988
 A:Title: The 3' portion of the gene for a Plasmodium yoelii merozoite surface antigen
 A:Reference number: A28121; MUID:88124889
 A:Accession: A28121
 A:Molecule type: DNA
 A:Residues: 1-680 <BUR>
 A:Cross-references: GB:J03612; NID:gl60678; PID:gl60679
 A:Experimental source: strain 17XL
 A:Note: the authors translated the codon GTA for residue 429 as Leu
 C:Superfamily: major merozoite surface antigen
 C:Keywords: surface antigen

Query Match 50.5%; Score 53; DB 2; Length 680;
 Best Local Similarity 55.6%; Pred. No. 2.4;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
 I:||||| : |||

DB 580 NAGCFRDNDGTGEWRCLL 597

RESULT 14

A45532
 A:Title: Cloning and analysis of the gene encoding the 230-kilodalton merozoite surface
 C:Species: Plasmodium yoelii
 C:Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 26-Aug-1999
 C:Accession: A45532; A45531
 R:Lewis, A.P.
 Mol. Biochem. Parasitol. 36, 271-282, 1989
 A:Title: Cloning and analysis of the gene encoding the 230-kilodalton merozoite surface
 A:Reference number: A45532; MUID:90014981
 A:Accession: A45532
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1772 <LEW>
 A:Cross-references: GB:J04668; NID:gl60492; PID:gl60493
 R:Daly, T.M.; Burns Jr., J.M.; Long, C.A.
 Mol. Biochem. Parasitol. 36, 283-285, 1989
 A:Title: Precursor to the major merozoite surface antigen of Plasmodium yoelii: cloning
 A:Reference number: A45531; MUID:90014982
 A:Accession: A45531
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 454-1094 <DAL>

A:Cross-references: GB:J03975; NID:gl60081; PID:gl60082
 C:Superfamily: major merozoite surface antigen
 C:Keywords: surface antigen

Query Match 50.5%; Score 53; DB 2; Length 1772;
 Best Local Similarity 55.6%; Pred. No. 5.5;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
 I:||||| : |||

DB 1672 NAGCFRDNDGTGEWRCLL 1689

RESULT 15

F82532
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
 C:Accession: F82532
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: F82532
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-298 <SIM>
 A:Cross-references: GB:AE004071; GB:AE003849; NID:g9107866; PID:AAF85442.1; GSPDB:GN
 A:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, L.E.A.; Carraro, D.M.; Carrer
 as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Perito, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; La
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
 , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
 M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
 A:Reference number: A59328
 A:Contents: annotation
 A:Genetics:
 A:Gene: XF2645
 C:Superfamily: conserved hypothetical protein HPI443

Query Match 43.8%; Score 46; DB 2; Length 298;
 Best Local Similarity 47.1%; Pred. No. 13;
 Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 SGCFRHLDERECKCLL 18
 I:||||| : |||

DB 253 SGCFFVEFSTRDEACAL 269

Search completed: January 29, 2002, 10:26:40
 Job time: 651 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:13:41 ; Search time 80.65 seconds
(without alignments)
8.183 Million cell updates/sec

Title: US-09-763-397a-11

Perfect score: 105

Sequence: 1 NSGCFRHLDERECKLL 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	100.0	1630	MSPI_PLAFK	P04932 plasmodium
2	105	100.0	1639	MSPI_PLAFW	P04933 plasmodium
3	105	100.0	1682	MSPI_PLAF3	P19598 plasmodium
4	105	100.0	1701	MSPI_PLAFF	P13819 plasmodium
5	105	100.0	1701	MSPI_PLAFM	P08569 plasmodium
6	105	100.0	1726	MSPI_PLAFC	P04934 plasmodium
7	105	100.0	1726	MSPI_PLAFP	P50495 plasmodium
8	53	50.5	1772	MSPI_PLAYO	P13828 plasmodium
9	46	43.8	298	ISPE_XYLFA	Q9pa75 xylella fas
10	46	43.8	537	KLC1_MOUSE	O88447 mus musculus
11	46	43.8	556	KLC1_RAT	P37285 rattus norv
12	46	43.8	569	KLC1_HUMAN	Q07866 homo sapien
13	45	42.9	451	CLUS_COTJA	P14018 coturnix co
14	45	42.9	2820	NFI_RAT	P97526 rattus norv
15	45	42.9	2839	NFI_HUMAN	P21359 homo sapien
16	45	42.9	2841	NFI_MOUSE	Q04690 mus musculus
17	44.5	42.4	116	HYPA_ECOLI	P24189 escherichia
18	44.5	42.4	2109	PGCA_CHICK	P07898 gallus gall
19	44	41.9	27	CKX7_CONPU	P56633 conus purpu
20	44	41.9	567	KDGE_HUMAN	P52429 homo sapien
21	43	41.0	796	PTPA_RAT	Q03348 rattus norv
22	43	41.0	802	PTPA_HUMAN	P18433 homo sapien
23	43	41.0	829	PTPA_MOUSE	P18052 mus musculus
24	42	40.0	75	YOR4_SMYEA	Q00848 strawberry
25	42	40.0	446	CLUS_PIG	Q29549 sus scrofa
26	41	39.0	387	PHEA_ERWHE	Q02286 erwinia her
27	41	39.0	439	CLUS_BOVIN	P17697 bos taurus
28	41	39.0	506	VAF1_SCHPO	Q09855 schizosacch
29	41	39.0	613	VQI6_CABEL	Q09279 caenorhabdi
30	41	39.0	882	LY10_HUMAN	Q13342 homo sapien
31	41	39.0	1109	MSH7_ARATH	Q9smv7 arabidopsis
32	40.5	38.6	1038	BMR2_HUMAN	Q13873 homo sapien
33	40.5	38.6	1038	BMR2_MOUSE	Q05607 mus musculus

34	40.5	38.6	3051	1	YNX3_CABEL	P34576 caenorhabdi
35	40	38.1	250	1	YC97_HUMAN	Q9y3a5 homo sapien
36	40	38.1	313	1	GAS2_HUMAN	O43903 homo sapien
37	40	38.1	314	1	GAS2_MOUSE	P11862 mus musculus
38	40	38.1	467	1	KIM6_HUMAN	O76013 homo sapien
39	40	38.1	777	1	UNC8_CABEL	Q1974 caenorhabdi
40	39.5	37.6	886	1	EMR1_HUMAN	Q14246 homo sapien
41	39	37.1	102	1	FTRC_GUITH	O78461 guillardia
42	39	37.1	108	1	RS23_ARATH	P49201 arabidopsis
43	39	37.1	111	1	FTRC_CYACA	Q9tm25 cyanidium c
44	39	37.1	118	1	FTRC_PORPU	P13886 porphyra pu
45	39	37.1	125	1	YDHL_ECOLI	P76188 escherichia

ALIGNMENTS

RESULT 1

ID	MSPI_PLAFK	STANDARD;	PRT;	1630 AA.
AC	P04932:			
DT	13-AUG-1987 (Rel. 05, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	MEZOZOITE SURFACE PROTEIN 1 PRECURSOR (MEZOZOITE SURFACE ANTIGENS) (PMMSA) (P190).			
DE	MSP-1.			
GN	Plasmodium falciparum (isolate K1 / Thailand).			
OS	Plasmodium falciparum (isolate K1 / Thailand).			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
OX	NCBI_TaxID=5839;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86136024; PubMed=3004972;			
RA	Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,			
RA	Stunnenberg H., Bujard H.;			
RT	*Polymorphism of the precursor for the major surface antigens of			
RT	Plasmodium falciparum merozoites: studies at the genetic level.*;			
RL	EMBO J. 4:3823-3829(1985).			
RN	[2]			
RP	REVISIONS, SEQUENCE FROM N.A.			
RA	Pan W., Tolle R., Bujard H.;			
RL	Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.			
CC	-1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (POTENTIAL).			
CC	-1- PFM: MEZOZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEZOZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X03371; CAA27070.1; -			
DR	PIR; A25120; SAQOK1.			
DR	InterPro; IPR000561; EGF-like.			
DR	Pfam; PF00008; EGF; 1.			
KW	Malaria; Merozoite; Polypeptide; GPI-anchored; Signal; Glycoprotein; Transmembrane; GPI-anchor.			
KW	SIGNAL	1	19	POTENTIAL.
FT	CHAIN	20	1630	MEZOZOITE SURFACE PROTEIN 1.
FT	DOMAIN	67	84	TRIPEPTIDE SG(TP) REPEAT.
FT	TRANSMEM	1614	1630	MEMBRANE ANCHOR.
FT	CARBOHYD	97	97	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	259	259	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	755	755	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	759	759	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	774	774	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	835	835	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	911	911	N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 955 955 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1156 1156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1436 1436 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1517 1517 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1630 AA; 187289 MW; ADBDEC3CE0A46322 CRC64;

Query Match 100.0%; Score 105; DB 1; Length 1630;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18
DB 1531 NSGCFRHLDERECKLL 1548

RESULT 2
MSPL_PLAFW STANDARD; PRT: 1639 AA.
AC P04933;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
DE (PMMSA) (P195).
GN MSP-1.
OS Plasmodium falciparum (isolate wellcome).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5848;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86014355; PubMed=2995820;
RA Holder A.A.; Lockyer M.J.; Odink K.G.; Sandhu J.S.; Riveros-Moreno V.;
RA Nicholas S.C.; Hillman Y.; Davey L.S.; Tizard M.L.V.; Schwarz R.T.;
RA Freeman R.R.;
RT "Primary structure of the precursor to the three major surface
RT antigens of Plasmodium falciparum merozoites.";
RL Nature 317:270-273(1985).
RN [2]
RP REVISIONS.
RA Holder A.A.;
RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
(POTENTIAL).
CC -!- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42
KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X02919; CAA26676.1; -.
CC FIRM; A24594; A24594.
CC InterPro: IPR000561; EGF-like.
CC Pfam: PF00008; EGF; 1.
CC Malaria; Merozoite; Polypeptide; Repeat; Signal; Glycoprotein;
KW Transmembrane, GPI-anchor.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1639 MEROZOITE SURFACE PROTEIN 1.
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 764 764 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 768 768 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 783 783 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 844 844 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 920 920 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 964 964 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1058 1058 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1174 1174 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1445 1445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1526 1526 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1639 AA; 187618 MW; 2C255B6616C87F6E CRC64;

Query Match 100.0%; Score 105; DB 1; Length 1639;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18
DB 1540 NSGCFRHLDERECKLL 1557

RESULT 3
MSPL_PLAF3 STANDARD; PRT: 1682 AA.
AC P19598; Q25921;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
DE (PMMSA) (P190).
GN MSP-1.
OS Plasmodium falciparum (isolate ro-33 / Ghana).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5834;
RN [1]
RP SEQUENCE OF 1-1061 FROM N.A.
RX MEDLINE=88166657; PubMed=3327688;
RA Certa U.; Rotman D.; Matile H.; Reber-Liske R.;
RT "A naturally occurring gene encoding the major surface antigen
RT precursor p190 of Plasmodium falciparum lacks tripeptide repeats.";
RL EMBO J. 6:4137-4142(1987).
RN [2]
RP SEQUENCE OF 1032-1682 FROM N.A.
RX MEDLINE=95354793; PubMed=7628566;
RA Tolle R.; Bujard H.; Cooper J.A.;
RT "Plasmodium falciparum: variations within the C-terminal region of
RT merozoite surface antigen-1.";
RL Exp. Parasitol. 81:47-54(1995).
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
(POTENTIAL).
CC -!- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42
KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M35727; AAA29715.1; -.
CC EMBL; Y00087; CAA68280.1; -.
CC EMBL; Z35326; CAA84555.1; -.
CC PIR; S06286; S06286.
CC InterPro: IPR000561; EGF-like.
CC Pfam; PF00008; EGF; 1.
CC Malaria; Merozoite; Polypeptide; Repeat; Signal; Glycoprotein;
KW Transmembrane, GPI-anchor.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1682 MEROZOITE SURFACE PROTEIN 1.
FT TRANSMEM 1666 1682 MEMBRANE ANCHOR (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 462 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 528 528 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 599 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 785 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 881 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 901 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 947 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1369 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1682 AA; 192462 MW; C82A1E159948CAD6 CRC64;

Query Match 100.0%; Score 105; DB 1; Length 1682;
 Best Local Similarity 100.0%; Pred. No. 1.8e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NSGCFRHLDERECKLL 18
 |||||
 DB 1583 NSGCFRHLDERECKLL 1600

RESULT 4
 MSPL_PLAFF
 ID MSPL_PLAFF STANDARD; PRT; 1701 AA.
 AC P13819;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
 DE (PMMSA).
 GN MSP-1.
 OS Plasmodium falciparum (isolate FC27 / Papua New Guinea).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5837;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88142999; PubMed=2449612;
 RA Peterson M.G., Coppel R.L., McIntyre P., Langford C.J., Woodrow G.,
 Brown G.V., Anders R.F., Kemp D.J.;
 RT "Variation in the precursor to the major merozoite surface antigens
 of Plasmodium falciparum.";
 RL Mol. Biochem. Parasitol. 27:291-302(1988).
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
 (POTENTIAL).
 CC -1- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42
 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
 MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M19143; AAA29653.1; .
 DR PIR; A54498; A54498.
 DR InterPro; IPR000561; EGF-like.
 DR Pfam; PF00008; EGF; 1.
 KW Malaria; Merozoite; Polypeptide; Repeat; Signal; Glycoprotein;
 KW Transmembrane; GPI-anchor.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 1701 MEROZOITE SURFACE PROTEIN 1.
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 470 470 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 802 802 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 899 899 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 919 919 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 965 965 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 991 991 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1089 1089 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1196 1196 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1588 1588 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1701 AA; 193719 MW; 3920875E73D38552 CRC64;

Query Match 100.0%; Score 105; DB 1; Length 1701;
 Best Local Similarity 100.0%; Pred. No. 1.8e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NSGCFRHLDERECKLL 18
 |||||
 DB 1602 NSGCFRHLDERECKLL 1619

RESULT 5
 MSPL_PLAFF
 ID MSPL_PLAFF STANDARD; PRT; 1701 AA.
 AC P08569;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
 DE (PMMSA) (P190).
 GN MSP-1.
 OS Plasmodium falciparum (isolate mad20 / Papua New Guinea).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=70153;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88011243; PubMed=3079521;
 RA Tanabe K., Mackay M., Goman M., Scaife J.G.;
 RT "Allelic dimorphism in a surface antigen gene of the malaria parasite
 Plasmodium falciparum.";
 RL J. Mol. Biol. 195:273-287(1987).
 CC [2]
 CC REVISIONS TO 1403; 1569 AND 1629.
 CC Tanabe K.;
 CC Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 CC [3]
 CC SEQUENCE OF 1-115 FROM N.A.
 CC MEDLINE=86136024; PubMed=3004972;
 CC Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,
 Stunnenberg H., Bulard H.;
 CC "Polymorphism of the precursor for the major surface antigens of
 Plasmodium falciparum merozoites: studies at the genetic level.";
 CC EMBO J. 4:3823-3829(1985).
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
 (POTENTIAL).
 CC -1- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42
 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
 MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X05624; CAA29112.1; .
 DR PIR; A26868; A26868.
 DR FIR; B25120; B25120.
 DR InterPro; IPR000561; EGF-like.
 DR Pfam; PF00008; EGF; 1.
 KW Malaria; Merozoite; Polypeptide; Repeat; Signal; Glycoprotein;
 KW Transmembrane; GPI-anchor.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 1701 MEROZOITE SURFACE PROTEIN 1.
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 470 470 N-LINKED (GLCNAC. . .) (POTENTIAL).


```
FT CARBOHYD 944 944 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 990 990 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1016 1016 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1114 1114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1221 1221 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1513 1513 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1726 AA; 196174 MW; 5B59CEEFA2F9A026 CRC64;

Query Match 100.0%; Score 105; DB 1; Length 1726;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSGCFRHLDERECKLL 18
|||||
Db 1627 NSGCFRHLDERECKLL 1644

RESULT 8
MSPI_PLAYO STANDARD; PRT; 1772 AA.
AC P13828;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
DE (PMISA) (230 KDA).
GN MSP-1.
OS Plasmodium berghei yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5862;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90205979; PubMed=2320061;
RA Lewis A.P.;
RT "Sequence analysis upstream of the gene encoding the precursor to the
major merozoite surface antigens of Plasmodium yoelii."
RL Mol. Biochem. Parasitol. 39:285-288(1990).
RN [2]
RP SEQUENCE OF 1093-1772 FROM N.A.
RC STRAIN=L7XL;
RX MEDLINE=86124889; PubMed=2448778;
RA Burns J.M. Jr., Daly T.M., Vaidya A.B., Long C.A.;
RT "The 3' portion of the gene for a Plasmodium yoelii merozoite surface
antigen encodes the epitope recognized by a protective monoclonal
antibody."
RL Proc. Natl. Acad. Sci. U.S.A. 85:602-606(1988).
CC -!- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42
KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; J03612; AAA29762.1; -.
DR EMBL; J04668; AAA29702.1; -.
DR PIR; A28121; A28121.
DR PIR; A45532; A45532.
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
Transmembrane; GPI-anchor.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 1772 MEROZOITE SURFACE PROTEIN 1.
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 646 646 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 829 829 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1090 1090 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1408 1408 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1446 1446 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1541 1541 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1629 1629 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1680 1680 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 1521 1521 L -> V (IN REF. 2).
SQ SEQUENCE 1772 AA; 197230 MW; 9A6291658E0F45D CRC64;

Query Match 50.5%; Score 53; DB 1; Length 1772;
Best Local Similarity 55.6%; Pred. No. 1.9;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 NSGCFRHLDERECKLL 18
|:|||||: ||:|
Db 1672 NAGCFRRDNGTEWRCLL 1689

RESULT 9
ISPE_XYLFA STANDARD; PRT; 298 AA.
AC Q9PA75;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE 4-DIPHOSPHOCYTIDYL-2-C-METHYL-D-ERYTHRITOL KINASE (EC 2.7.1.-) (CMK)
DE [4-(CYTIDINE-5'-DIPHOSPHO)-2-C-METHYL-D-ERYTHRITOL KINASE].
GN ISPE OR XF2645.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OX Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.D., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshuko M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa."
RL Nature 406:151-159(2000).
CC -!- FUNCTION: CATALYZES THE PHOSPHORYLATION OF THE POSITION 2 HYDROXY
GROUP OF 4-DIPHOSPHOCYTIDYL-2C-METHYL-D-ERYTHRITOL (BY
SIMILARITY)
CC -!- PATHWAY: DEOXYXYLULOSE-5-PHOSPHATE PATHWAY (DXP) OF ISOPRENOID
BIOSYNTHESIS; FOURTH STEP.
CC -!- SIMILARITY: BELONGS TO THE ISPE FAMILY.
-----
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AE004071; AAF85442.1; -
 KW Transferase; Kinase; Isoprene biosynthesis; ATP-binding;
 KW Complete proteome. 119 ATP (POTENTIAL).
 FT NP_BIND 109
 SQ SEQUENCE 298 AA; 31800 MW; 972DB2DAF4722AD1 CRC64;

Query Match 43.8%; Score 46; DB 1; Length 298;
 Best Local Similarity 47.1%; Pred. No. 4;
 Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 SGCFRHLDERECKLL 18
 Db 253 SGCFVFSTRDEACAL 269

RESULT 10
 KLC1_MOUSE STANDARD; PRT; 537 AA.
 AC O88447;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE KINESIN LIGHT CHAIN 1 (KLC 1).
 DE KNS2 OR KLC1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C; TISSUE=Brain;
 RX MEDLINE=98288268; PubMed=9624122;
 RA Rahman A., Friedman D.S., Goldstein L.S.;
 RT "Two kinesin light chain genes in mice. Identification and
 RT characterization of the encoded proteins.";
 RL J. Biol. Chem. 273:15395-15403(1998).
 CC -!- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING
 CC PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT. THE LIGHT
 CC CHAIN MAY FUNCTION IN COUPLING OF CARGO TO THE HEAVY CHAIN OR IN
 CC THE MODULATION OF ITS ATPASE ACTIVITY.
 CC -!- SUBUNIT: OLIGOMERIC COMPLEX COMPOSED OF TWO HEAVY CHAINS AND
 CC TWO LIGHT CHAINS.
 CC -!- SIMILARITY: BELONGS TO THE KINESIN LIGHT CHAIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 6 TPR REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF055655; AAC27740.1; ALT_INIT.
 DR MGD; MGI:107978; Klc1.
 DR InterPro; IPR002151; Kinesin_light.
 DR InterPro; IPR001440; TPR.
 DR Pfam; PF00515; TPR; 5.
 DR PRINTS; PR00381; KINESINLIGHT.
 DR SMART; SM00028; TPR; 4.
 DR PROSITE; PS01160; KINESIN_LIGHT; 3.
 KW Motor protein; Microtubules; Coiled coil; Repeat; TPR repeat.
 FT DOMAIN 23 152
 FT REPEAT 207 240 TPR 1.

FT REPEAT 249 282 TPR 2.
 FT REPEAT 291 324 TPR 3.
 FT REPEAT 333 366 TPR 4.
 FT REPEAT 376 409 TPR 5.
 FT REPEAT 459 492 TPR 6.
 SQ SEQUENCE 537 AA; 60926 MW; 1EE2E52444F287B5 CRC64;

Query Match 43.8%; Score 46; DB 1; Length 537;
 Best Local Similarity 53.3%; Pred. No. 7.1;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECK 15
 Db 418 NKPIWMHAERECK 432

RESULT 11
 KLC1_RAT STANDARD; PRT; 556 AA.
 AC P37285;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE KINESIN LIGHT CHAIN 1 (KLC 1).
 DE KNS2 OR KLC1 OR KLC.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=92052221; PubMed=1946431;
 RA Cyr J.L., Pfister K.K., Bloom G.S., Slaughter C.A., Brady S.T.;
 RT "Molecular genetics of kinesin light chains: generation of isoforms
 RT by alternative splicing.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:10114-10118(1991).
 CC -!- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING
 CC PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT. THE LIGHT
 CC CHAIN MAY FUNCTION IN COUPLING OF CARGO TO THE HEAVY CHAIN OR IN
 CC THE MODULATION OF ITS ATPASE ACTIVITY.
 CC -!- SUBUNIT: OLIGOMERIC COMPLEX COMPOSED OF TWO HEAVY CHAINS AND
 CC TWO LIGHT CHAINS.
 CC -!- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; A, B AND C (SHOWN
 CC HERE); ARE PRODUCED BY ALTERNATIVE SPLICING. THEY ONLY DIFFER IN
 CC THE C-TERMINUS.
 CC -!- SIMILARITY: BELONGS TO THE KINESIN LIGHT CHAIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 6 TPR REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M75148; -; NOT_ANNOTATED_CDS.
 DR EMBL; M75147; -; NOT_ANNOTATED_CDS.
 DR EMBL; M75146; -; NOT_ANNOTATED_CDS.
 DR InterPro; IPR002151; Kinesin_light.
 DR InterPro; IPR001440; TPR.
 DR Pfam; PF00515; TPR; 5.
 DR PRINTS; PR00381; KINESINLIGHT.
 DR SMART; SM00028; TPR; 4.
 DR PROSITE; PS01160; KINESIN_LIGHT; 4.
 KW Motor protein; Microtubules; Coiled coil; Repeat; TPR repeat;
 FT DOMAIN 23 152
 FT REPEAT 209 242 COILED COIL.
 FT REPEAT 251 284 TPR 1.
 FT REPEAT 293 326 TPR 2.
 FT REPEAT 326 TPR 3.


```
FT REPEAT      335 368 TPR 4.
FT REPEAT      377 410 TPR 5.
FT REPEAT      460 493 TPR 6.
FT VARSPLIC    538 546 MISSING (IN ISOFORM B).
FT VARSPLIC    538 556 VMSVSEWNGMRKMKGLVK -> A (IN ISOFORM A).
SQ SEQUENCE    556 AA; 63247 MW; 1A843B34BCDDB53E CRC64;

Query Match      43.8%; Score 46; DB 1; Length 556;
Best Local Similarity 53.3%; Pred. No. 7.3;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECK 15
DB 419 NKPIWMHAERECK 433

RESULT 12
ID KLC1_HUMAN STANDARD; PRT; 569 AA.
AC Q07866;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE KINESIN LIGHT CHAIN 1 (KLC 1).
GN KNS2 OR KLC1 OR KLC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94099888; PubMed=8274221;
RA Cabeza-Arvelaiz Y., Shih L.-C.N., Hardman N., Asselbergs F.,
RA Bilbe G., Schmitz A., White B., Siciliano M.J., Lachman L.B.;
RT "Cloning and genetic characterization of the human kinesin light-chain
RL DNA Cell Biol. 12:881-892(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Kaplan J.;
RA Gerber S.; Rozet J.M., Perrault I., Ducroq D., Souled E., Munnich A.,
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING
CC PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT. THE LIGHT
CC CHAIN MAY FUNCTION IN COUPLING OF CARGO TO THE HEAVY CHAIN OR IN
CC THE MODULATION OF ITS ATPASE ACTIVITY.
CC -!- SUBUNIT: OLIGOMERIC COMPLEX COMPOSED OF TWO HEAVY CHAINS AND
CC TWO LIGHT CHAINS.
CC -!- ALTERNATIVE PRODUCTS: AS FOR THE RAT KINESIN LIGHT CHAIN, THE
CC HUMAN PRODUCT MAY WELL UNDERGO ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: FOUND IN A VARIETY OF TISSUES. MOSTLY ABUNDANT
CC IN BRAIN AND SPINE.
CC -!- SIMILARITY: BELONGS TO THE KINESIN LIGHT CHAIN FAMILY.
CC -!- SIMILARITY: CONTAINS 6 TPR REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L04733; AAL16576.1; -
DR EMBL; AF267530; AAF72543.1; -
DR EMBL; AF267518; AAF72543.1; JOINED.
DR EMBL; AF267519; AAF72543.1; JOINED.
DR EMBL; AF267520; AAF72543.1; JOINED.
DR EMBL; AF267521; AAF72543.1; JOINED.
DR EMBL; AF267522; AAF72543.1; JOINED.
DR EMBL; AF267523; AAF72543.1; JOINED.
DR EMBL; AF267524; AAF72543.1; JOINED.
```

```
DR EMBL; AF267525; AAF72543.1; JOINED.
DR EMBL; AF267526; AAF72543.1; JOINED.
DR EMBL; AF267527; AAF72543.1; JOINED.
DR EMBL; AF267528; AAF72543.1; JOINED.
DR EMBL; AF267529; AAF72543.1; JOINED.
DR MIM; 600025; -
DR InterPro; IPR002151; Kinesin_light.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00515; TPR 5.
DR PRINTS; PR00381; KINESINLIGHT.
DR SMART; SM00028; TPR 4.
DR PROSITE; PS01160; KINESIN_LIGHT; 4.
KW Motor protein; Microtubules; Coiled coil; Repeat; TPR repeat;
KW Alternative splicing.
FT DOMAIN      23 152 COILED COIL.
FT REPEAT      209 242 TPR 1.
FT REPEAT      251 284 TPR 2.
FT REPEAT      293 326 TPR 3.
FT REPEAT      335 368 TPR 4.
FT REPEAT      377 410 TPR 5.
FT REPEAT      460 493 TPR 6.
SQ SEQUENCE    569 AA; 64786 MW; 5D9376C3EEDD00FE CRC64;

Query Match      43.8%; Score 46; DB 1; Length 569;
Best Local Similarity 53.3%; Pred. No. 7.5;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECK 15
DB 419 NKPIWMHAERECK 433

RESULT 13
ID CLUS_COTJA STANDARD; PRT; 451 AA.
AC P14018;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CLUSTERIN PRECURSOR (51.5 KDA PROTEIN).
GN T64.
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Neuroretina;
RX MEDLINE=89239492; PubMed=2541393;
RA Michel D., Gillet G., Volovitch M., Pessac B., Calothy G., Brun G.;
RT "Expression of a novel gene encoding a 51.5 kD precursor protein is
RT induced by different retroviral oncogenes in quail neuroretinal
RT cells."
RL Oncogene Res. 4:127-136(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95262670; PubMed=7744033;
RA Michel D., Chatelein G., Herault Y., Brun G.;
RT "The expression of the avian clusterin gene can be driven by two
RT alternative promoters with distinct regulatory elements."
RL Eur. J. Biochem. 229:215-223(1995).
CC -!- FUNCTION: NOT YET CLEAR. IT IS KNOWN TO BE EXPRESSED IN A VARIETY
CC OF TISSUES AND IT SEEMS TO BE ABLE TO BIND TO CELLS, MEMBRANES,
CC AND HYDROPHOBIC PROTEINS. IT HAS BEEN ASSOCIATED WITH PROGRAMMED
CC CELL DEATH.
CC -!- SUBUNIT: ANTIPARALLEL DISULFIDE-LINKED HETERODIMER
CC (BY SIMILARITY).
CC -!- INDUCTION: BY DIFFERENT RETROVIRAL ONCOGENES.
CC -!- SIMILARITY: BELONGS TO THE CLUSTERIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X15825; CAA33823.1; -
EMBL; X60760; CAA56733.1; -
PIR; S07714; S07714.
InterPro; IPR000753; Clusterin.
Pfam; PF01093; Clusterin; 1.
SMART; SM00035; Cla; 1.
SMART; SM00030; Clb; 1.
PROSITE; PS00492; CLUSTERIN_1; 1.
PROSITE; PS00493; CLUSTERIN_2; 1.
KW Glycoprotein; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 451 CLUSTERIN.
FT CHAIN 19 230
FT CHAIN 231 451
FT CHAIN 231 451
FT DISULFID 98 314 BETA-CHAIN (SMALL SUBUNIT) (POTENTIAL).
FT DISULFID 109 306 ALPHA-CHAIN (LARGE SUBUNIT) (POTENTIAL).
FT DISULFID 112 303 INTERCHAIN (BY SIMILARITY).
FT DISULFID 117 296 INTERCHAIN (BY SIMILARITY).
FT DISULFID 125 286 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 375 375 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 437 437 A -> P (IN REF. 1).
FT CONFLICT 445 451 KONNTIE -> SRTTP (IN REF. 1).
SQ SEQUENCE 451 AA; 51800 MW; 8CF81125C5B22F3 CRC64;

Query Match 42.9%; Score 45; DB 1; Length 451;
Best Local Similarity 44.4%; Pred. No. 8.5;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Oy 1 NSGCFRHLDERECKCLL 18
::||| ||| |::|:
Db 293 SAGCLRMRDECEKREIL 310

RESULT 14
NF1_RAT
ID NF1_RAT STANDARD; PRT; 2820 AA.
AC P97526; 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NEUROFIBROMIN (NEUROFIBROMATOSIS-RELATED PROTEIN NF-1).
GN NF1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RA Kyrtis A.P., Lee P.S., Mochizuki H., Nishi T., Levin V.A., Saya H.;
RT "Differential splicing of the neurofibromatosis type 1 (NF1) gene in
RT rats: Homologous splice variants in human are expressed in rat
RT cells.";
RL Int. J. Oncol. 1:149-152(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RA MEDLINE=97137538; PubMed=8982875;
RX Suzuki H., Takahashi K., Yasumoto K., Fuse N., Shibahara S.;
RT "Differential tissue-specific expression of neurofibromin isoform
RT mRNAs in rat.";

J. Biochem. 120:1048-1054(1996).
-!- FUNCTION: STIMULATES THE GTPASE ACTIVITY OF RAS. NF1 SHOWS GREATER
CC AFFINITY FOR RAS GAP, BUT LOWER SPECIFIC ACTIVITY. MAY BE A
CC REGULATOR OF RAS ACTIVITY (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 RAS-GAP DOMAIN.
CC

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; D45201; BAA08141.1; -
DR InterPro; IPR001251; CRAL_TRIO.
DR InterPro; IPR001936; RasGAP.
DR Pfam; PF00616; RasGAP; 1.
DR SMART; SM00323; RasGAP; 1.
DR SMART; SM00516; SEC14; 1.
DR PROSITE; PS00509; RAS_GTPASE_ACTIV_1; 1.
DR PROSITE; PS00018; RAS_GTPASE_ACTIV_2; 1.
KW GTPase activation.
FT DOMAIN 1237 1432 RAS-GAP.
FT DOMAIN 1354 1357 POLY-SER.
SQ SEQUENCE 2820 AA; 317080 MW; FCL08487E86DA89F CRC64;

Query Match 42.9%; Score 45; DB 1; Length 2820;
Best Local Similarity 53.8%; Pred. No. 50;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 4 CFRHLDERECKC 16
||||| |::|:
Db 711 CFRHLEEDIRC 723

RESULT 15
NF1_HUMAN

ID NF1_HUMAN STANDARD; PRT; 2839 AA.
AC P21359; Q00662; Q14284; Q14930; Q9UMK3;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NEUROFIBROMIN (NEUROFIBROMATOSIS-RELATED PROTEIN NF-1).
GN NF1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=93090270; PubMed=1457041;
RA Bernards A., Haase V.H., Murthy A.E., Menon A., Hannigan G.E.,
RA Gusella J.F.;
RT "Complete human NF1 cDNA sequence: two alternatively spliced mRNAs
RT and absence of expression in a neuroblastoma line.";
RL DNA Cell Biol. 11:727-734(1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=90319792; PubMed=2134734;
RA Wallace M.R., Marchuk D.A., Andersen L.B., Letcher R., Odeh H.M.,
RA Saulino A.M., Fountain J.W., Brereton A., Nicholson J., Mitchell A.L.,
RA Brownstein B.H., Collins F.S.;
RT "Type 1 neurofibromatosis gene: identification of a large transcript
RT disrupted in three NF1 patients.";
RL Science 249:181-186(1990).
RN [3]
RP ERRATUM.
RX MEDLINE=91102559; PubMed=2125369;
RA Wallace M.R., Marchuk D.A., Andersen L.B., Collins F.S.;
RL Science 250:1749-1749(1990).
RN [4]

RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE-92147138; PubMed-1783401;
RA Marchuk D.A., Saulino A.M., Tavakkol R., Swaroop M., Wallace M.R.,
RA Andersen L.B., Mitchell A.L., Gutmann D.H., Boguski M.S.,
RA Collins F.S.;
RT "cDNA cloning of the type 1 neurofibromatosis gene: complete sequence
of the NF1 gene product.";
RL Genomics 11:931-940(1991).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RX MEDLINE-92412152; PubMed-1339276;
RA Suzuki H., Takahashi K., Kubota Y., Shibahara S.;
RT "Molecular cloning of a cDNA coding for neurofibromatosis type 1
protein isoform lacking the domain related to ras GTPase-activating
protein.";
RL Biochem. Biophys. Res. Commun. 187:984-990(1992).
RN [6]
RP SEQUENCE OF 335-2839 FROM N.A. (ISOFORM 1), AND VARIANT PRO-1953.
RX MEDLINE-90304909; PubMed-2114220;
RA Cawthon R.M., Weiss R., Xu G., Viskochil D., Culver M., Stevens J.,
RA Robertson M., Dunn D., Gesteland R., O'Connell P., White R.;
RT "A major segment of the neurofibromatosis type 1 gene: cDNA sequence,
genomic structure, and point mutations.";
RL Cell 62:193-201(1990).
RN [7]
RP SEQUENCE OF 335-2839 FROM N.A. (ISOFORM 1).
RX MEDLINE-90335969; PubMed-2116237;
RA Xu G., O'Connell P., Viskochil D., Cawthon R.M., Robertson M.,
RA Culver M., Dunn D., Stevens J., Gesteland R., White R., Weiss R.;
RT "The neurofibromatosis type 1 gene encodes a protein related to GAP.";
RL Cell 62:599-608(1990).
RN [8]
RP SEQUENCE OF 707-782 FROM N.A.
RX MEDLINE-97156204; PubMed-9002664;
RA Regnier V., Meddeb M., Lecointre G., Richard F., Duverger A.,
RA Nguyen V.C., Dutrillaux B., Bernheim A., Danglot G.;
RT "Emergence and scattering of multiple neurofibromatosis (NF1)-related
sequences during hominoid evolution suggest a process of
pericentromeric interchromosomal transposition.";
RL Hum. Mol. Genet. 6:9-16(1997).
RN [9]
RP SEQUENCE OF 751-1611 FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE-95293414; PubMed-7774960;
RA Li Y., O'Connell P., Breidenbach H.H., Cawthon R.M., Stevens J.,
RA Xu G., Neil S., Robertson M., White R., Viskochil D.;
RT "Genomic organization of the neurofibromatosis 1 gene (NF1).";
RL Genomics 25:9-18(1995).
RN [10]
RP SEQUENCE OF 1090-1598 FROM N.A. (ISOFORM 4).
RX MEDLINE-91029515; PubMed-2121370;
RA Martin G.A., Viskochil D., Bollag G., McCabe P.C., Crosier W.J.,
RA Haubruck H., Controy L., Clark R., O'Connell P., Cawthon R.M.,
RA Innis M., McCormick F.;
RT "The GAP-related domain of the neurofibromatosis type 1 gene product
interacts with ras p21.";
RL Cell 63:843-849(1990).
RN [11]
RP SEQUENCE OF 1168-1566 FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE-92019823; PubMed-1923522;
RA Nishi T., Lee P.S., Oka K., Levin V.A., Tanase S., Morino Y.,
RA Saya H.;
RT "Differential expression of two types of the neurofibromatosis type 1
(NF1) gene transcripts related to neuronal differentiation.";
RL Oncogene 6:1555-1559(1991).
RN [12]
RP SEQUENCE OF 1371-1391 FROM N.A. (ISOFORM 2).
RX MEDLINE-93109335; PubMed-8417346;
RA Andersen L.B., Ballester R., Marchuk D.A., Chang E., Gutmann D.H.,
RA Saulino A.M., Canonis J., Wigler M., Collins F.S.;
RT "A conserved alternative splice in the von Recklinghausen
neurofibromatosis (NF1) gene produces two neurofibromin isoforms,
both of which have GTPase-activating protein activity.";
RL Mol. Cell. Biol. 13:487-495(1993).
RN [13]
RP SEQUENCE OF 1371-1391 FROM N.A. (ISOFORM 2).
RX MEDLINE-92109786; PubMed-1662505;
RA Suzuki Y., Suzuki H., Kayama T., Yoshimoto T., Shibahara S.;
RT "Brain tumors predominantly express the neurofibromatosis type 1
gene transcripts containing the 63 base insert in the region coding
for GTPase activating protein-related domain.";
RL Biochem. Biophys. Res. Commun. 181:955-961(1991).
RN [14]
RP FUNCTION
RX MEDLINE-91029516; PubMed-2121371;
RA Ballester R., Marchuk D.A., Boguski M.S., Saulino A.M., Letcher R.,
RA Wigler M., Collins F.S.;
RT "The NF1 locus encodes a protein functionally related to mammalian
GAP and yeast IRA proteins.";
RL Cell 63:851-859(1990).
RN [15]
RP REVIEW ON VARIANTS.
RX MEDLINE-95072625; PubMed-7981724;
RA Upadhyaya M., Shaw D.J., Harper P.S.;
RT "Molecular basis of neurofibromatosis type 1 (NF1): mutation analysis
and polymorphisms in the NF1 gene.";
RL Hum. Mutat. 4:83-101(1994).
RN [16]
RP REVIEW ON VARIANTS.
RX MEDLINE-96422425; PubMed-8825042;
RA Shen M.H., Harper P.S., Upadhyaya M.;
RT "Molecular genetics of neurofibromatosis type 1 (NF1).";
RL J. Med. Genet. 33:2-17(1996).
RN [17]
RP VARIANT GLU-1444.
RX MEDLINE-92233464; PubMed-1568247;
RA Li Y., Bollag G., Clark R., Stevens J., Controy L., Ward K.,
RA Friedman E., Samowitz W., Robertson M., Bradley P., McCormick F.,
RA White R., Cawthon R.M.;
RT "Somatic mutations in the neurofibromatosis 1 gene in human tumors.";
RL Cell 69:275-281(1992).
RN [18]
RP VARIANTS MET-2164 AND ASN-2192.
RX MEDLINE-93258316; PubMed-1302608;
RA Upadhyaya M., Shen M.H., Cherryson A., Farnham J., Maynard J.,
RA Huson S.M., Harper P.S.;
RT "Analysis of mutations at the neurofibromatosis 1 (NF1) locus.";
RL Hum. Mol. Genet. 1:735-740(1992).
RN [19]
RP VARIANT HIS-1721--LEU-1733 DUPL.
RX MEDLINE-93304433; PubMed-8317503;
RA Tassabehji M., Strachan T., Sharland M., Colley A., Donnai D.,
RA Harris R., Thakker N.;
RT "Tandem duplication within a neurofibromatosis type 1 (NF1) gene exon
in a family with features of Watson syndrome and Noonan syndrome.";
RL Am. J. Hum. Genet. 53:90-95(1993).
RN [20]
RP VARIANT MET-991 DEL.
RX MEDLINE-94108439; PubMed-7904209;
RA Shen M.H., Harper P.S., Upadhyaya M.;
RT "Neurofibromatosis type 1 (NF1): the search for mutations by PCR-
heteroduplex analysis on Hydrolink gels.";
RL Hum. Mol. Genet. 2:1861-1864(1993).
RN [21]
RP VARIANTS NF1 ASP-1166 AND ARG-1440.
RX MEDLINE-95072575; PubMed-7981679;
RA Purandare S.M., Lanyon W.G., Connor J.M.;
RT "Characterisation of inherited and sporadic mutations in
neurofibromatosis type-1.";
RL Hum. Mol. Genet. 3:1109-1115(1994).
RN [22]
RP VARIANT NF1 2387-ASN-PHE-2388 DEL.
RX MEDLINE-94362704; PubMed-8081387;
RA Abernathy C.R., Colman S.D., Kousseff B.G., Wallace M.R.;
RT "Two NF1 mutations: frameshift in the GAP-related domain, and loss of
two codons toward the 3' end of the gene.";
RL Hum. Mutat. 3:347-352(1994).

RN [23]
 RP VARIANT NF1 ALA-2631.
 RX MEDLINE=96091873; PubMed=8544190;
 RA Upadhyaya M., Maynard J., Osborn M.J., Huson S.M., Ponder M.,
 RA Ponder B.A.J., Harper P.S.;
 RT "Characterisation of germline mutations in the neurofibromatosis type
 RT 1 (NF1) gene";
 RL J. Med. Genet. 32:706-710(1995).
 RN [24]
 RP VARIANT NF1 ARG-629.
 RX MEDLINE=96431167; PubMed=8834249;
 RA Gasparini P., D'Agruma L., de Cillis G.P., Balestrazzi P.,
 RA Mingarelli R., Zelante L.;
 RT "Scanning the first part of the neurofibromatosis type 1 gene by RNA-
 RT SSCP: identification of three novel mutations and of two new
 RT polymorphisms.";
 RL Hum. Genet. 97:492-495(1996).
 RN [25]
 RP VARIANT NF1 ARG-1035.
 RX MEDLINE=96400960; PubMed=8807336;

Query Match 42.9%; Score 45; DB 1; Length 2839;
 Best Local Similarity 53.8%; Pred. NO. 51;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 CFRHLDERECKC 16
 DB 709 CFRHLCCEADIRC 721

Search completed: January 29, 2002, 11:13:42
 Job time: 818 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:12:12 ; Search time 285.36 Seconds
(without alignments)
9.227 Million cell updates/sec

Title: US-09-763-397A-11
Perfect score: 105
Sequence: 1 NSGCFRHLDERECKLL 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	100.0	76	5 Q9BJS8	Q9BJS8 plasmodium
2	105	100.0	77	5 Q9BJV6	Q9BJV6 plasmodium
3	105	100.0	77	5 Q9BJV5	Q9BJV5 plasmodium
4	105	100.0	77	5 Q9BJV4	Q9BJV4 plasmodium
5	105	100.0	77	5 Q9BJV3	Q9BJV3 plasmodium
6	105	100.0	77	5 Q9BJV2	Q9BJV2 plasmodium
7	105	100.0	77	5 Q9BJV1	Q9BJV1 plasmodium
8	105	100.0	77	5 Q9BJV0	Q9BJV0 plasmodium
9	105	100.0	77	5 Q9BJU9	Q9BJU9 plasmodium
10	105	100.0	77	5 Q9BJU8	Q9BJU8 plasmodium
11	105	100.0	77	5 Q9BJU7	Q9BJU7 plasmodium
12	105	100.0	77	5 Q9BJU6	Q9BJU6 plasmodium
13	105	100.0	77	5 Q9BJU5	Q9BJU5 plasmodium
14	105	100.0	77	5 Q9BJU4	Q9BJU4 plasmodium
15	105	100.0	77	5 Q9BJU3	Q9BJU3 plasmodium
16	105	100.0	77	5 Q9BJU2	Q9BJU2 plasmodium
17	105	100.0	77	5 Q9BJU1	Q9BJU1 plasmodium
18	105	100.0	77	5 Q9BJU0	Q9BJU0 plasmodium
19	105	100.0	77	5 Q9BJT9	Q9BJT9 plasmodium

20	105	100.0	77	5 Q9BJT8	Q9BJT8 plasmodium
21	105	100.0	77	5 Q9BJT7	Q9BJT7 plasmodium
22	105	100.0	77	5 Q9BJT6	Q9BJT6 plasmodium
23	105	100.0	77	5 Q9BJT5	Q9BJT5 plasmodium
24	105	100.0	77	5 Q9BJT4	Q9BJT4 plasmodium
25	105	100.0	77	5 Q9BJT3	Q9BJT3 plasmodium
26	105	100.0	77	5 Q9BJT2	Q9BJT2 plasmodium
27	105	100.0	77	5 Q9BJT1	Q9BJT1 plasmodium
28	105	100.0	77	5 Q9BJT0	Q9BJT0 plasmodium
29	105	100.0	77	5 Q9BJS9	Q9BJS9 plasmodium
30	105	100.0	77	5 Q9BJS7	Q9BJS7 plasmodium
31	105	100.0	77	5 Q9BJS6	Q9BJS6 plasmodium
32	105	100.0	96	5 Q9TYE5	Q9TYE5 plasmodium
33	105	100.0	116	5 Q9TYE3	Q9TYE3 plasmodium
34	105	100.0	119	5 Q9TYE7	Q9TYE7 plasmodium
35	105	100.0	121	5 Q9UA18	Q9UA18 plasmodium
36	105	100.0	121	5 Q9UA17	Q9UA17 plasmodium
37	105	100.0	121	5 Q9UA16	Q9UA16 plasmodium
38	105	100.0	121	5 Q9TYE6	Q9TYE6 plasmodium
39	105	100.0	121	5 Q9TYE4	Q9TYE4 plasmodium
40	105	100.0	218	5 Q9TVG8	Q9TVG8 plasmodium
41	105	100.0	219	5 Q9UB87	Q9UB87 plasmodium
42	105	100.0	372	5 Q43997	Q43997 plasmodium
43	105	100.0	372	5 Q25717	Q25717 plasmodium
44	105	100.0	372	5 Q25718	Q25718 plasmodium
45	105	100.0	372	5 Q25725	Q25725 plasmodium

ALIGNMENTS

RESULT 1
Q9BJS8 ID Q9BJS8 PRELIMINARY; PRT; 76 AA.
AC Q9BJS8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID:5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface protein 1-19 and merozoite surface protein 2 in independent field isolates of Plasmodium falciparum."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329535; AAK19354.1; -
FT NON_TER 1
FT NON_TER 76
SQ SEQUENCE 76 AA; 8406 MW; EC87FA0F5F3C903F CRC64;

Query Match 100.0%; Score 105; DB 5; Length 76;
Best Local Similarity 100.0%; Pred. No. 4.8e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18
Db 4 NSGCFRHLDERECKLL 21

RESULT 2
Q9BJV6 ID Q9BJV6 PRELIMINARY; PRT; 77 AA.
AC Q9BJV6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.

```
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329507; AAK19326.1; -
FT NON_TER 1
FT NON_TER 77
FT SEQUENCE 77 AA; 8542 MW; 4EBC8D8B8AA169D4 CRC64;

Query Match 100.0%; Score 105; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 4.8e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
Db 4 NSGCFRHLDERECKCLL 21

RESULT 3
ID Q9BJV5 PRELIMINARY; PRT; 77 AA.
AC Q9BJV5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329508; AAK19327.1; -
FT NON_TER 1
FT NON_TER 77
FT SEQUENCE 77 AA; 8542 MW; 4EBC8D8B8AA169D4 CRC64;

Query Match 100.0%; Score 105; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 4.8e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
Db 4 NSGCFRHLDERECKCLL 21

RESULT 4
ID Q9BJV4 PRELIMINARY; PRT; 77 AA.
AC Q9BJV4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
```

```
RT isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329509; AAK19328.1; -
FT NON_TER 1
FT NON_TER 77
FT SEQUENCE 77 AA; 8542 MW; 4EBC8D8B8AA169D4 CRC64;

Query Match 100.0%; Score 105; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 4.8e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
Db 4 NSGCFRHLDERECKCLL 21

RESULT 5
ID Q9BJV3 PRELIMINARY; PRT; 77 AA.
AC Q9BJV3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329510; AAK19329.1; -
FT NON_TER 1
FT NON_TER 77
FT SEQUENCE 77 AA; 8542 MW; 4EBC8D8B8AA169D4 CRC64;

Query Match 100.0%; Score 105; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 4.8e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
Db 4 NSGCFRHLDERECKCLL 21

RESULT 6
ID Q9BJV2 PRELIMINARY; PRT; 77 AA.
AC Q9BJV2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329511; AAK19330.1; -
FT NON_TER 1
FT NON_TER 77
FT SEQUENCE 77 AA; 8542 MW; 4EBC8D8B8AA169D4 CRC64;
```

Query Match 100.0%; Score 105; DB 5; Length 77;
 Best Local Similarity 100.0%; Pred. No. 4.8e-10;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
 ID 4 NSGCFRHLDERECKCLL 21
 DB 4 NSGCFRHLDERECKCLL 21

RESULT 7
 Q9BJV1 PRELIMINARY; PRT; 77 AA.
 AC Q9BJV1;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
 RT "Antigenic drift and immune selection acting on merozoite surface
 RT protein 1-19 and merozoite surface protein 2 in independent field
 RT isolates of Plasmodium falciparum."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF329513; AAK19331.1; -
 FT NON_TER 1
 FT NON_TER 77
 SQ SEQUENCE 77 AA; 8542 MW; 4EBC8D8B8AA169D4 CRC64;

Query Match 100.0%; Score 105; DB 5; Length 77;
 Best Local Similarity 100.0%; Pred. No. 4.8e-10;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
 ID 4 NSGCFRHLDERECKCLL 21
 DB 4 NSGCFRHLDERECKCLL 21

RESULT 8
 Q9BJV0 PRELIMINARY; PRT; 77 AA.
 AC Q9BJV0;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
 RT "Antigenic drift and immune selection acting on merozoite surface
 RT protein 1-19 and merozoite surface protein 2 in independent field
 RT isolates of Plasmodium falciparum."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF329513; AAK19332.1; -
 FT NON_TER 1
 FT NON_TER 77
 SQ SEQUENCE 77 AA; 8542 MW; 4EBC8D8B8AA169D4 CRC64;

Query Match 100.0%; Score 105; DB 5; Length 77;
 Best Local Similarity 100.0%; Pred. No. 4.8e-10;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
 ID 4 NSGCFRHLDERECKCLL 21
 DB 4 NSGCFRHLDERECKCLL 21

Db 4 NSGCFRHLDERECKCLL 21

RESULT 9
 Q9BJU9 PRELIMINARY; PRT; 77 AA.
 AC Q9BJU9;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
 RT "Antigenic drift and immune selection acting on merozoite surface
 RT protein 1-19 and merozoite surface protein 2 in independent field
 RT isolates of Plasmodium falciparum."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF329514; AAK19333.1; -
 FT NON_TER 1
 FT NON_TER 77
 SQ SEQUENCE 77 AA; 8542 MW; 4EBC8D8B8AA169D4 CRC64;

Query Match 100.0%; Score 105; DB 5; Length 77;
 Best Local Similarity 100.0%; Pred. No. 4.8e-10;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
 ID 4 NSGCFRHLDERECKCLL 21
 DB 4 NSGCFRHLDERECKCLL 21

RESULT 10
 Q9BJU8 PRELIMINARY; PRT; 77 AA.
 AC Q9BJU8;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
 RT "Antigenic drift and immune selection acting on merozoite surface
 RT protein 1-19 and merozoite surface protein 2 in independent field
 RT isolates of Plasmodium falciparum."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF329515; AAK19334.1; -
 FT NON_TER 1
 FT NON_TER 77
 SQ SEQUENCE 77 AA; 8614 MW; 4EA799E1BAA169D4 CRC64;

Query Match 100.0%; Score 105; DB 5; Length 77;
 Best Local Similarity 100.0%; Pred. No. 4.8e-10;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18
 ID 4 NSGCFRHLDERECKCLL 21
 DB 4 NSGCFRHLDERECKCLL 21

RESULT 11
 Q9BJU7 PRELIMINARY; PRT; 77 AA.
 AC Q9BJU7;
 ID Q9BJU7;
 DB 4 NSGCFRHLDERECKCLL 21

DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE MERZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
ON NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
protein 1-19 and merozoite surface protein 2 in independent field
isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF329516; AAK19335.1; -
FT NON_TER 1 1
FT NON_TER 77 77
SQ SEQUENCE 77 AA; 8614 MW; 4EA799F4B4E69D4 CRC64;
[1]
Query Match 100.0%; Score 105; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 4.8e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NSGCFRHLDERECKCLL 18
Db 4 NSGCFRHLDERECKCLL 21
[1]
RESULT 12
Q9BJU6
ID Q9BJU6 PRELIMINARY; PRT; 77 AA.
AC Q9BJU6;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE MERZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
ON NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
protein 1-19 and merozoite surface protein 2 in independent field
isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF329517; AAK19336.1; -
FT NON_TER 1 1
FT NON_TER 77 77
SQ SEQUENCE 77 AA; 8488 MW; 4EBC984F4B4E69D4 CRC64;
[1]
Query Match 100.0%; Score 105; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 4.8e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NSGCFRHLDERECKCLL 18
Db 4 NSGCFRHLDERECKCLL 21
[1]
RESULT 13
Q9BJU5
ID Q9BJU5 PRELIMINARY; PRT; 77 AA.
AC Q9BJU5;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE MERZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
ON NCBI_TaxID=5833;

RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
protein 1-19 and merozoite surface protein 2 in independent field
isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF329518; AAK19337.1; -
FT NON_TER 1 1
FT NON_TER 77 77
SQ SEQUENCE 77 AA; 8587 MW; 4EA799FF4B4E69D4 CRC64;
[1]
Query Match 100.0%; Score 105; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 4.8e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NSGCFRHLDERECKCLL 18
Db 4 NSGCFRHLDERECKCLL 21
[1]
RESULT 14
Q9BJU4
ID Q9BJU4 PRELIMINARY; PRT; 77 AA.
AC Q9BJU4;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE MERZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
ON NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
protein 1-19 and merozoite surface protein 2 in independent field
isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF329519; AAK19338.1; -
FT NON_TER 1 1
FT NON_TER 77 77
SQ SEQUENCE 77 AA; 8587 MW; 4EA799FF4B4E69D4 CRC64;
[1]
Query Match 100.0%; Score 105; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 4.8e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NSGCFRHLDERECKCLL 18
Db 4 NSGCFRHLDERECKCLL 21
[1]
RESULT 15
Q9BJU3
ID Q9BJU3 PRELIMINARY; PRT; 77 AA.
AC Q9BJU3;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE MERZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
ON NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
protein 1-19 and merozoite surface protein 2 in independent field
isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF329520; AAK19339.1; -
FT NON_TER 1
FT NON_TER 77
SQ SEQUENCE 77 AA; 8587 MW; 4EA799FF4B4E69D4 CRC64;

Query Match 100.0%; Score 105; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 4.8e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSGCFRHLDERECKLL 18
Db 4 NSGCFRHLDERECKLL 21

Search completed: January 29, 2002, 11:12:12
Job time: 768 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:21:45 ; Search time 310.82 Seconds
(without alignments)
4.528 Million cell updates/sec

Title: US-09-763-397A-12

Perfect score: 106

Sequence: 1 EDGSGNGKITCECTKPD 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq.ll01:*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	106	100.0	19	21	Plasmodium falciparum
2	106	100.0	54	14	MSPIEGF2A EGF2-lik
3	106	100.0	96	22	Merozoite surface
4	106	100.0	108	22	Merozoite surface
5	106	100.0	116	18	pfMSPI(p19)A prote
6	106	100.0	116	18	pfMSPI(p19)A prote
7	106	100.0	127	18	pfMSPI(p19)S prote
8	106	100.0	127	18	pfMSPI(p19)S prote
9	106	100.0	350	21	Recombinant vaccin
10	106	100.0	355	20	Merozoite surface
11	106	100.0	355	20	Merozoite surface

12	106	100.0	361	20	AAV09373
13	106	100.0	361	20	AAV05833
14	106	100.0	376	20	AAV09374
15	106	100.0	376	20	AAV05834
16	106	100.0	1639	19	AAV54145
17	102	96.2	1654	6	AAV50777
18	93	87.7	54	14	AAV41357
19	93	87.7	375	22	AAV83926
20	48	45.3	160	21	AAV54107
21	47	44.3	106	14	AAV41358
22	44	41.5	66	16	AAV09392
23	44	41.5	66	16	AAV62970
24	43.5	41.0	3329	21	AAV56029
25	43	40.6	66	15	AAV57134
26	43	40.6	66	16	AAV09390
27	43	40.6	66	16	AAV09391
28	43	40.6	66	16	AAV62968
29	43	40.6	66	16	AAV62969
30	43	40.6	84	16	AAV09389
31	43	40.6	84	16	AAV62967
32	43	40.6	518	22	AAV92474
33	43	40.6	740	22	AAE01030
34	42.5	40.1	130	21	AAE12499
35	42.5	40.1	150	21	AAE12497
36	42.5	40.1	171	22	AAE68362
37	42	39.6	93	22	AAV37611
38	42	39.6	97	22	AAV32889
39	42	39.6	100	21	AAV01957
40	42	39.6	100	21	AAV01960
41	42	39.6	104	14	AAV38596
42	42	39.6	104	19	AAV58496
43	42	39.6	109	22	AAV69602
44	42	39.6	110	12	AAV13183
45	42	39.6	110	18	AAV27548

ALIGNMENTS

RESULT 1

AAV70288

ID AAV70288 standard; peptide; 19 AA.

XX AAV70288;

XX AAV70288;

DT 06-JUN-2000 (first entry)

DE Plasmodium falciparum MSP-1 antigenic epitope, P598.

XX Recombinant protein; CDC/NIIMVAC-1; multivalent; malaria; vaccine;
KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
KW Circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
KW Liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
KW Apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
KW Pfg27; antiparasitic; prevention; anti-CDC/NIIMVAC-1 antibody.

OS Plasmodium falciparum.

XX WO200011179-A1.

PN 02-MAR-2000.

XX 19-AUG-1999; 99WO-US18869.

XX 21-AUG-1998; 98US-0097703.

PA (NAIM-) NAT INST IMMUNOLOGY.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Lal AA, Shi YP, Hasnain SE;

XX WPI; 2000-237654/20.

XX Novel recombinant protein as vaccine for treating malarial infection
 PT comprises antigenic peptides obtained from different stages of
 PT plasmodium falciparum life cycle -
 XX
 PS Claim 2; Page 16; 52pp; English.
 XX
 CC The present sequence is the antigenic epitope p598, derived from
 CC merozoite surface protein-1 (MSP-1) of the asexual blood stage of
 CC Plasmodium falciparum. It is used in the construction of recombinant
 CC vaccine CDC/N1MALVAC-1, which is a multivalent, multistage malarial
 CC vaccine. The recombinant protein comprises, melittin signal peptide,
 CC (His)6 tag, T-cell epitope from tetanus toxoid and 21 antigenic epitopes
 CC from circumsporozoite protein (CSP), sporozoite surface protein-2
 CC (SSP-2), liver stage antigen-1 (LSA-1), merozoite surface protein-1
 CC (MSP-1), MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding
 CC antigen-175 (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete
 CC specific antigen, Pf27. These epitopes were obtained at different stages
 CC of the life cycle of P. falciparum. CDC/N1MALVAC-1 vaccine has
 CC antiparasitic activity and can be used for treatment and prevention of
 CC malarial infections. Anti-CDC/N1MALVAC-1 antibodies can be used for
 CC detecting P. falciparum in biological samples.
 XX
 SQ Sequence 19 AA;

Query Match 100.0%; Score 106; DB 21; Length 19;
 Best Local Similarity 100.0%; Pred. No. 8.1e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPDs 19
 |||||
 Db 1 edsgsngkkitcectkpdS 19

RESULT 2
 AAR41356
 ID AAR41356 standard; peptide; 54 AA.

XX AAR41356;
 XX
 DT 04-MAR-1994 (first entry)
 XX
 DE MSP1EGF2A EGF2-like domain variant.

XX Epidermal growth factor 1; merozoite surface protein 1; malaria;
 KW vaccine.
 XX

OS Plasmodium yoelii.

XX Key Location/Qualifiers
 FH Cleavage-site 1
 FT /note= "introduced to facilitate cleavage
 FT from recombinant protein"

XX WO9317107-A.
 XX
 PD 02-SEP-1993.
 XX
 PF 22-FEB-1993; 93WO-GB00367.
 XX
 PR 22-FEB-1992; 92GB-0003821.
 XX
 XX (MEDI-) MEDICAL RES COUNCIL.
 XX
 PI Blackman MJ, Chappel JA, Holder AA;
 XX
 DR WPI; 1993-288413/36.

XX Allelic variants of epidermal growth factor 1- or 2-like domains - of
 PT merozoite surface protein 1, produced recombinantly for malaria
 PT vaccines
 XX

PS Claim 2; Fig 2a; 35pp; English.

XX The sequence is that of an allelic variant of a merozoite surface
 CC protein-1 epidermal growth factor (EGF) 2-like domain. It may be
 CC used alone or as part of a fusion protein of EGF-1-like and
 CC EGF-2-like domains in vaccines against malaria. When expressed
 CC recombinantly it is produced in a form indistinguishable from
 CC that in the native protein.
 XX
 SQ Sequence 54 AA;

Query Match 100.0%; Score 106; DB 14; Length 54;
 Best Local Similarity 100.0%; Pred. No. 2.2e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPDs 19
 |||||
 Db 23 edsgsngkkitcectkpdS 41

RESULT 3
 AAB37608
 ID AAB37608 standard; protein; 96 AA.

XX AAB37608;
 XX
 DT 27-FEB-2001 (first entry)
 XX

DE Merozoite surface protein-1.

XX Merozoite surface protein; protazoacide; vaccine; malaria.

XX Plasmodium falciparum.

XX WO200063245-A2.

XX 26-OCT-2000.

XX 20-APR-2000; 2000WO-GB01558.

XX 20-APR-1999; 99GB-0009072.

XX 13-MAY-1999; 99US-0311817.

XX 23-MAY-1999; 99CA-2271451.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Holder A, Birdsall B, Reeney J, Morgan W, Syed S, Uthaipibull C;

XX WPI; 2001-015762/02.

XX Novel variants of the C-terminal fragment of Plasmodium merozoite
 PT surface protein-1, useful as vaccines for treating or preventing
 PT malaria -

XX Example 2; Page 48; 126pp; English.

XX The present invention relates to non-natural variants of a C-terminal
 CC fragment of a Plasmodium merozoite surface protein-1 (MSP-1). The
 CC non-natural variants have reduced affinity for at least 1 antibody
 CC capable of blocking a second antibody that inhibits the proteolytic
 CC cleavage of Plasmodium MSP-1_4_2, and has the same affinity for at least
 CC one third antibody that inhibits the proteolytic cleavage of Plasmodium
 CC MSP-1_4_2, compared to natural MSP-1_1_9. The present sequence is the
 CC wild-type MSP-1 protein. This sequence was used to generate the variants
 CC of the present invention. The non-natural variants of the present
 CC invention are useful for immunising a mammal against malaria, and can be
 CC used to treat malaria.

XX Sequence 96 AA;

Query Match 100.0%; Score 106; DB 22; Length 96;

Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPDs 19
DB 65 edsgsngkkitcectkpd 83

RESULT 4
AA37609
ID AAB37609 standard; Protein; 108 AA.
XX
AC AAB37609;
XX
DT 27-FEB-2001 (first entry)
XX
DE Merozoite surface protein-119.
XX
KW Merozoite surface protein; protazoacide; vaccine; malaria.
XX
OS Plasmodium falciparum.
XX
PN WO200063245-A2.
XX
PD 26-OCT-2000.
XX
PF 20-APR-2000; 2000WO-GB01558.
XX
PR 20-APR-1999; 99GB-0009072.
PR 13-MAY-1999; 99US-0311817.
PR 25-MAY-1999; 99CA-2271451.
XX
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Holder A, Birdsall B, Feeney J, Morgan W, Syed S, Uthaipibull C;
XX
DR WPI; 2001-015762/02.
DR N-PSDB; AAC68977.
XX
PT Novel variants of the C-terminal fragment of Plasmodium merozoite
PT surface protein-1, useful as vaccines for treating or preventing
PT malaria -
XX
PS Example 5; Fig 15; 126pp; English.
XX
CC The present invention relates to non-natural variants of a C-terminal
CC fragment of a Plasmodium merozoite surface protein-1 (MSP-1). The
CC non-natural variants have reduced affinity for at least 1 antibody
CC capable of blocking a second antibody that inhibits the proteolytic
CC cleavage of Plasmodium MSP-1.4.2, and has the same affinity for at least
CC one third antibody that inhibits the proteolytic cleavage of Plasmodium
CC MSP-1.4.2, compared to natural MSP-1.1.9. The non-natural variants of the
CC present invention are useful for immunising a mammal against malaria, and
CC can be used to treat malaria. The present sequence is MSP-119 protein.
XX
SQ Sequence 108 AA;

Query Match 100.0%; Score 106; DB 22; Length 108;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPDs 19
DB 77 edsgsngkkitcectkpd 95

RESULT 5
AA36103
ID AA36103 standard; Protein; 116 AA.
XX
AC AA36103;
XX

DT 25-MAR-1998 (first entry)
XX
DE PfMSP1(p19)A protein sequence.
XX
KW Plasmodium vivax; merozoite surface protein; MSP1; p19;
KW Plasmodium falciparum; malaria; vaccine; immunity; epitope.
XX
OS Plasmodium falciparum.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT 1..95 /note= "amino acids derived from P. falciparum MSP1 p19
FT fragment"
FT Region 96..116
FT /note= "glycosylphosphatidylinositol anchoring sequence"
XX
PN WO9730158-A2.
XX
PD 21-AUG-1997.
XX
PF 14-FEB-1997; 97WO-FR00290.
XX
PR 14-FEB-1996; 96FR-0001822.
XX
PA (INSP) INST PASTEUR.
PA (UYNK) UNIV NEW YORK STATE.
XX
PI Barnwell JW, Longacre-Andre S, Mendis K, Nato F;
PI Roth C;
XX
DR WPI; 1997-425033/39.
DR N-PSDB; AAT94550.
XX
PT Recombinant protein containing the merozoite surface protein-1 p19
PT fragment - useful in anti-malarial vaccines, diagnosis and protein
PT purification
XX
PS Disclosure; Fig 1B; 85pp; French.
XX
CC This is the amino acid sequence of a recombinant protein comprising
CC amino acids 1613-1705 of the Plasmodium falciparum merozoite surface
CC protein 1 (MSP1) 19 kD C-terminal fragment (p19), linked to a
CC glycosylphosphatidylinositol membrane anchoring sequence. p19 is the
CC C-terminal fragment of the 42 kD MSP1 from Plasmodium species.
CC The recombinant protein can be used for the production of anti-malarial
CC vaccines, where the p19 fragment provides a high level of protective
CC immunity since it includes epitopes not presented in the p42 fragment.
XX
SQ Sequence 116 AA;

Query Match 100.0%; Score 106; DB 18; Length 116;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPDs 19
DB 67 edsgsngkkitcectkpd 85

RESULT 6
AA22592
ID AA22592 standard; Protein; 116 AA.
XX
AC AA22592;
XX
DT 25-MAR-1998 (first entry)
XX
DE PfMSP1(p19)A protein sequence.
XX
KW Plasmodium vivax; merozoite surface protein; MSP1; p19;
KW Plasmodium falciparum; malaria; vaccine; immunity; epitope.

```

XX OS Plasmodium falciparum.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Region 1..95
XX FT /note= "amino acids derived from P. falciparum MSP1 p19
XX FT fragment"
XX FT Region 96..116
XX FT /note= "glycosylphosphatidylinositol anchoring sequence"
XX FT WO9730159-A2.
XX PN 21-AUG-1997.
XX XX
XX PD 14-FEB-1997; 97WO-FR00291.
XX PF 14-FEB-1996; 96FR-0001821.
XX PR
XX PA (INSP ) INST PASTEUR.
XX PA (UYN ) UNIV NEW YORK STATE.
XX PI Barnwell JW, Longacre-Andre S, Mendis K, Nato F;
XX PI Roth C;
XX DR WPI; 1997-425034/39.
XX DR P-PSDB; AAW22592.
XX XX
XX PT Recombinant protein containing Plasmodium merozoite surface
XX PT protein-1 p42 fragment - useful in antimalarial vaccines, also new
XX PT antibodies for diagnosis and protein purification
XX XX
XX PS Disclosure; Fig 1B; 85pp; French.
XX XX
XX CC This is the amino acid sequence of a recombinant protein comprising
XX CC amino acids 1613-1705 of the Plasmodium falciparum merozoite surface
XX CC protein 1 (MSP1) 19 kD C-terminal fragment (p19), linked to a
XX CC glycosylphosphatidylinositol membrane anchoring sequence. p19 is the
XX CC C-terminal fragment of the 42 kD MSP1 from Plasmodium species.
XX CC The recombinant protein can be used for the production of anti-malarial
XX CC vaccines, where the p19 fragment provides a high level of protective
XX CC immunity since it includes epitopes not presented in the p42 fragment.
XX CC
XX SQ Sequence 116 AA;

Query Match 100.0%; Score 106; DB 18; Length 116;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPD5 19
Db 67 edsgsngkkitcectkpd5 85

RESULT 7
AAW22593
ID AAW22593 standard; Protein; 127 AA.
XX AC
XX AC AAW22593;
XX DT 25-MAR-1998 (first entry)
XX DE PFMSPI(p19)S protein sequence.
XX XX
XX KW Chimeric; Plasmodium vivax; merozoite surface protein; MSP1; p19;
XX KW Plasmodium falciparum; malaria; vaccine; immunity; epitope.
XX XX
XX OS Chimeric - Plasmodium vivax.
XX OS Chimeric - Plasmodium falciparum.
XX FH Key Location/Qualifiers
XX FT Peptide 1..19

```

```

FT FT Protein /note= "signal peptide"
FT FT 20..127 /note= "mature protein"
FT FT Region 1..32
FT FT /note= "derived from P. vivax MSP1"
FT FT 33..34
FT FT /note= "encoded by restriction enzyme sequence used to
FT FT create the chimeric sequence"
FT FT Region 35..127
FT FT /note= "derived from P. falciparum C-terminal p19
FT FT fragment of MSP1"
FT XX
XX PN WO9730159-A2.
XX XX
XX PD 21-AUG-1997.
XX XX
XX PF 14-FEB-1997; 97WO-FR00291.
XX XX
XX PR 14-FEB-1996; 96FR-0001821.
XX XX
XX PA (INSP ) INST PASTEUR.
XX PA (UYN ) UNIV NEW YORK STATE.
XX XX
XX PI Barnwell JW, Longacre-Andre S, Mendis K, Nato F;
XX PI Roth C;
XX XX
XX DR WPI; 1997-425034/39.
XX DR P-PSDB; AAW22592.
XX XX
XX PT Recombinant protein containing Plasmodium merozoite surface
XX PT protein-1 p42 fragment - useful in antimalarial vaccines, also new
XX PT antibodies for diagnosis and protein purification
XX XX
XX PS Disclosure; Fig 1C; 85pp; French.
XX XX
XX CC This is the amino acid sequence of a chimeric protein comprising amino
XX CC acids 1-32 of the Plasmodium vivax merozoite surface protein 1 (MSP1)
XX CC linked to amino acids Asn1613-Ser1705 which correspond to the 19 kD
XX CC C-terminal fragment (p19) of MSP-1 from Plasmodium falciparum.
XX CC p19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species.
XX CC The recombinant protein can be used for the production of anti-malarial
XX CC vaccines, where the p19 fragment provides a high level of protective
XX CC immunity since it includes epitopes not presented in the p42 fragment.
XX CC
XX SQ Sequence 127 AA;

Query Match 100.0%; Score 106; DB 18; Length 127;
Best Local Similarity 100.0%; Pred. No. 4.8e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPD5 19
Db 99 edsgsngkkitcectkpd5 117

RESULT 8
AAW36102
ID AAW36102 standard; Protein; 127 AA.
XX AC
XX AC AAW36102;
XX DT 25-MAR-1998 (first entry)
XX DE PFMSPI(p19)S protein sequence.
XX XX
XX KW Chimeric; Plasmodium vivax; merozoite surface protein; MSP1; p19;
XX KW Plasmodium falciparum; malaria; vaccine; immunity; epitope.
XX XX
XX OS Chimeric - Plasmodium vivax.
XX OS Chimeric - Plasmodium falciparum.
XX FH Key Location/Qualifiers

```

```
FT Peptide 1..19
FT /note= "signal peptide"
FT Protein 20..127
FT /note= "mature protein"
FT Region 1..32
FT /note= "derived from P. vivax MSP1"
FT Region 33..34
FT /note= "encoded by restriction enzyme sequence used to
FT create the chimeric sequence"
FT Region 35..127
FT /note= "derived from P. falciparum C-terminal p19
FT fragment of MSP1"
XX
PN W09730158-A2.
XX
XX 21-AUG-1997.
XX
XX 14-FEB-1997; 97WO-FR00290.
XX
XX 14-FEB-1996; 96FR-0001822.
XX
XX (INSP ) INST PASTEUR.
XX (UYNY ) UNIV NEW YORK STATE.
XX
XX Barnwell JW, Longacre-Andre S, Mendis K, Nato F;
XX Roth C;
XX
XX WPI: 1997-425033/39.
XX N-PSDB; AAT94549.
XX
XX Recombinant protein containing the merozoite surface protein-1 p19
XX fragment - useful in anti-malarial vaccines, diagnosis and protein
XX purification
XX
XX Disclosure; Fig 1C; 85pp; French.
XX
XX This is the amino acid sequence of a chimeric protein comprising amino
XX acids 1-32 of the Plasmodium vivax merozoite surface protein 1 (MSP1)
XX linked to amino acids Asn1613-Ser1705 which correspond to the 19 kD
XX C-terminal fragment (p19) of MSP-1 from Plasmodium falciparum.
XX
XX P19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species.
XX The recombinant protein can be used for the production of anti-malarial
XX vaccines, where the p19 fragment provides a high level of protective
XX immunity since it includes epitopes not presented in the p42 fragment.
XX
XX Sequence 127 AA;

Query Match 100.0%; Score 106; DB 18; Length 127;
Best Local Similarity 100.0%; Pred. No. 4.8e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPDs 19
Db 99 edsgsngkkitcectkpd 117
|||||
RESULT 9
RAY70278
ID AAY70278 standard; Protein; 350 AA.
XX
XX AAY70278;
XX
XX 06-JUN-2000 (first entry)
XX
XX Recombinant vaccine CDC/N1MALVAC-1.
XX
XX Recombinant protein; CDC/N1MALVAC-1; multivalent; malaria; vaccine;
XX T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
XX circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
XX liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
XX apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
XX EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
```

```
KW Pfg27; antiparasitic; prevention; anti-CDC/N1MALVAC-1 antibody;
KW honey bee.
XX
XX Chimeric - Apis sp.
OS Chimeric - Clostridium tetani.
OS Chimeric - Plasmodium falciparum.
XX
XX Key Location/Qualifiers
FT Peptide 1..22
FT /label= Melittin_signal_peptide
FT /note= "Derived from Honey bee"
FT Protein 23..350
FT /label= Mature_CDC/N1MALVAC-1
FT /note= "Recombinant multivalent malarial vaccine"
XX
XX W0200011179-A1.
XX
XX 02-MAR-2000.
XX
XX 19-AUG-1999; 99WO-US18869.
XX
XX 21-AUG-1998; 98US-0097703.
XX
XX (NAIM-) NAT INST IMMUNOLOGY.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Lal AA, Shi YP, Hasnain SE;
XX
XX WPI: 2000-237654/20.
XX N-PSDB; AAZ51336.
XX
XX Novel recombinant protein as vaccine for treating malarial infection
XX comprises antigenic peptides obtained from different stages of
XX plasmodium falciparum life cycle -
XX
XX Claim 3; Page 43-44; 52pp; English.
XX
XX The present sequence is that of recombinant protein CDC/N1MALVAC-1,
XX which is a multivalent, multistage malarial vaccine. The recombinant
XX protein comprises, melittin signal peptide, (His)6 tag, T-cell epitope
XX from tetanus toxoid and 21 antigenic epitopes from circumsporozoite
XX protein (CSP), sporozoite surface protein-2 (SSP-2), liver stage
XX antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1), MSP-2, apical
XX membrane antigen-1 (AMA-1), erythrocyte binding antigen-175 (EBA-175),
XX rhoptry associated protein-1 (RAP-1) and gamete specific antigen, Pfg27.
XX These epitopes were obtained at different stages of the life cycle of
XX Plasmodium falciparum. CDC/N1MALVAC-1 vaccine has antiparasitic
XX activity and can be used for treatment and prevention of malarial
XX infections. Anti-CDC/N1MALVAC-1 antibodies can be used for detecting
XX P. falciparum in biological samples.
XX
XX Sequence 350 AA;

Query Match 100.0%; Score 106; DB 21; Length 350;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPDs 19
Db 100 edsgsngkkitcectkpd 118
|||||
RESULT 10
RAY09372
ID AAY09372 standard; Protein; 355 AA.
XX
XX AAY09372;
XX
XX 31-AUG-1999 (first entry)
XX
XX Merozoite surface protein MSP-1-42.
XX
```


XX 15-MAY-1998; 98US-0085649.
 PR 20-OCT-1997; 97US-0062592.
 XX (GENZ) GENZYME TRANSGENICS CORP.
 PA Chen LH, Meade H;
 PI WPI: 1999-288313/24.
 XX P-PSDB; AAX56009.
 DR Modified malarial protein for use in anti-malarial vaccines
 XX Example; Fig 2; 35pp; English.
 XX This present sequence comprises a 42 kDa C-terminal portion of
 CC malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an
 CC important target for the development of a vaccine against
 CC Plasmodium falciparum. The C-terminal end of the sequence is
 CC modified to include a 6xHis tag. A nucleic acid (see AAX56008)
 CC encoding MSP-1-42 has been modified according to a method
 CC of the invention in order to improve expression in mammalian cells
 CC and in transgenic animals by reducing the AT content and removing
 CC mRNA instability motifs. The invention allows expression of
 CC MSP-1-42 in the milk of transgenic animals, and also provides a DNA
 CC vaccine comprising a vector containing the altered MSP-1-42 nucleic
 CC acid.
 XX Sequence 361 AA;
 SQ Query Match 100.0%; Score 106; DB 20; Length 361;
 Best Local Similarity 100.0%; Pred. No. 1.3e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCCTKPDs 19
 |||||
 DB 327 edsgsgngkkitcctkpd 345

RESULT 13
 AAY05833
 ID AAY05833 standard; Protein; 361 AA.
 XX AC AAY05833;
 XX DT 02-AUG-1999 (first entry)
 XX DE Merozoite surface protein MSP-1-42.
 KW MSP-1; merozoite surface protein; malaria; vaccine;
 KW protein engineering; protein expression; codon usage;
 KW transgenic animal.
 XX Plasmodium falciparum.
 OS Synthetic.
 XX Key
 FH Peptide 1..15 Location/Qualifiers
 FT /note= "beta-casein signal peptide"
 FT Peptide 371..376
 FT /note= "6xHis tag"
 FT Misc-difference 197
 FT /note= "Asn in native MSP-1-42 (N181Q mutation)"
 FT Misc-difference 278
 FT /note= "Asn in native MSP-1-42 (N262Q mutation)"
 XX WO9920774-A2.
 XX 29-APR-1999.
 XX 20-OCT-1998; 98WO-US22225.
 XX 15-MAY-1998; 98US-0085649.
 PR 20-OCT-1997; 97US-0062592.
 XX (GENZ) GENZYME TRANSGENICS CORP.
 PA Chen LH, Meade H;
 PI WPI: 1999-302742/25.
 XX N-PSDB; AAX25587.
 XX New modified recombinant nucleic acid sequences useful for producing

PT malarial DNA vaccine
 XX Disclosure; Fig 2; 43pp; English.
 XX This present sequence comprises a 42 kDa C-terminal portion of
 CC malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an
 CC important target for the development of a vaccine against
 CC Plasmodium falciparum. The C-terminal end of the sequence is
 CC modified to include a 6xHis tag. Nucleic acids (see AAX25586 and
 CC AAX25593) encoding MSP-1-42 have been modified according to a method
 CC of the invention in order to improve expression in mammalian cells
 CC and in transgenic animals. The invention provides modified
 CC recombinant nucleic acid sequences and methods for increasing the
 CC mRNA levels and protein expression of proteins that are difficult
 CC to express in cell culture systems, especially mammalian cell
 CC culture systems or in transgenic animals. The preferred difficult
 CC protein candidates for expression are those derived from lower
 CC organisms such as parasites, bacteria and viruses that have DNA
 CC coding sequences of high AT content or which have mRNA instability
 CC motifs or rare codons relative to the recombinant expression system
 CC to be used. The invention allows expression of MSP-1 in the milk
 CC of transgenic animals, and also provides a DNA vaccine comprising a
 CC vector containing the altered MSP-1-42 nucleic acid.
 XX Sequence 361 AA;
 SQ Query Match 100.0%; Score 106; DB 20; Length 361;
 Best Local Similarity 100.0%; Pred. No. 1.3e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCCTKPDs 19
 |||||
 DB 327 edsgsgngkkitcctkpd 345

RESULT 14
 AAY09374
 ID AAY09374 standard; Protein; 376 AA.
 XX AC AAY09374;
 XX DT 31-AUG-1999 (first entry)
 XX DE Modified merozoite surface protein MSP-1-42.
 KW MSP-1; merozoite surface protein; malaria; vaccine;
 KW protein engineering; protein expression; codon usage;
 KW transgenic animal; mutant.
 XX Plasmodium falciparum.
 OS Synthetic.
 XX Key
 FH Peptide 1..15 Location/Qualifiers
 FT /note= "beta-casein signal peptide"
 FT Peptide 371..376
 FT /note= "6xHis tag"
 FT Misc-difference 197
 FT /note= "Asn in native MSP-1-42 (N181Q mutation)"
 FT Misc-difference 278
 FT /note= "Asn in native MSP-1-42 (N262Q mutation)"
 XX WO9920774-A2.
 XX 29-APR-1999.
 XX 20-OCT-1998; 98WO-US22226.
 XX 15-MAY-1998; 98US-0085649.
 PR 20-OCT-1997; 97US-0062592.
 XX (GENZ) GENZYME TRANSGENICS CORP.
 PA

```

XX  PI      Chen LH, Meade H;
XX  WPI; 1999-288313/24.
XX  DR      P-PSDB; AAX56008.
XX  XX
XX  PT      Modified malarial protein for use in anti-malarial vaccines
XX  PS      Example; Fig 11; 35pp; English.
XX  CC      The present sequence represents a 42 kDa C-terminal portion of
XX  CC      malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an
XX  CC      important target for the development of a vaccine against
XX  CC      Plasmodium falciparum. The sequence has been modified to include
XX  CC      an N-terminal beta-casein signal peptide and an N-terminal 6xHis
XX  CC      tag. In addition, N181Q and N262Q mutations have been introduced to
XX  CC      eliminate 2 N-glycosylation sites. These modifications allow the
XX  CC      MSP-1-42 protein to be expressed in the milk of transgenic mice.
XX  SQ      Sequence 376 AA;
XX
XX  Query Match 100.0%; Score 106; DB 20; Length 376;
XX  Best Local Similarity 100.0%; Pred. No. 1.3e-06;
XX  Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX  Qy 1 EDGSGNGKKITCECTKPDs 19
XX  Db 342 edsgsngkkitcectkpd 360
XX
XX  RESULT 15
XX  ID      AAY05834
XX  AC      AAY05834;
XX  DT      02-AUG-1999 (first entry)
XX  DE      Modified merozoite surface protein MSP-1-42.
XX  KW      MSP-1; merozoite surface protein; malaria; vaccine;
XX  KW      protein engineering; protein expression; codon usage;
XX  KW      transgenic animal; mutant.
XX  OS      Plasmodium falciparum.
XX  OS      Synthetic.
XX  FH      Key Location/Qualifiers
XX  FT      Peptide 1..15
XX  FT      /note= "beta-casein signal peptide"
XX  FT      Peptide 371..376
XX  FT      /note= "6xHis tag"
XX  FT      Misc-difference 197
XX  FT      /note= "Asn in native MSP-1-42 (N181Q mutation)"
XX  FT      Misc-difference 278
XX  FT      /note= "Asn in native MSP-1-42 (N262Q mutation)"
XX  PN      WO9920766-A2.
XX  XX
XX  PD      29-APR-1999.
XX  PF      20-OCT-1998; 98WO-US22225.
XX  PR      15-MAY-1998; 98US-0085649.
XX  PR      20-OCT-1997; 97US-0062592.
XX  PA      (GENZ ) GENZYME TRANSGENICS CORP.
XX  PI      Chen LH, Meade H;
XX  WPI; 1999-302742/25.
XX  DR      N-PSDB; AAX25593.

```

```

XX  PT      New modified recombinant nucleic acid sequences useful for producing
XX  XX      malarial DNA vaccine
XX  PS      Example; Fig 11; 43pp; English.
XX  CC      The present sequence represents a 42 kDa C-terminal portion of
XX  CC      malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an
XX  CC      important target for the development of a vaccine against
XX  CC      Plasmodium falciparum. The sequence has been modified to include
XX  CC      an N-terminal beta-casein signal peptide and an N-terminal 6xHis
XX  CC      tag. In addition, N181Q and N262Q mutations have been introduced to
XX  CC      eliminate 2 N-glycosylation sites. These modifications allow the
XX  CC      MSP-1-42 protein to be expressed in the milk of transgenic mice.
XX  CC      The invention generally relates to modified recombinant nucleic
XX  CC      acid sequences and methods for increasing the mRNA levels and
XX  CC      protein expression of proteins that are difficult to express in
XX  CC      cell culture systems, mammalian cell culture systems or in
XX  CC      transgenic animals. The preferred difficult protein candidates for
XX  CC      expression are those derived from lower organisms such as parasites,
XX  CC      bacteria and viruses that have DNA coding sequences of high AT
XX  CC      content or which have mRNA instability motifs or rare codons
XX  CC      relative to the recombinant expression system to be used.
XX  SQ      Sequence 376 AA;
XX
XX  Query Match 100.0%; Score 106; DB 20; Length 376;
XX  Best Local Similarity 100.0%; Pred. No. 1.3e-06;
XX  Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX  Qy 1 EDGSGNGKKITCECTKPDs 19
XX  Db 342 edsgsngkkitcectkpd 360
XX

```

Search completed: January 29, 2002, 10:21:46
Job time: 422 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:24:05 ; Search time 133.18 Seconds
(without alignments)
3.210 Million cell updates/sec

Title: US-09-763-397A-12

Perfect score: 106
Sequence: 1 EDGSGNGKITCECTKPD5 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCrUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	106	100.0	53	1	US-08-290-919-3
2	93	87.7	53	1	US-08-290-919-4
3	47	44.3	106	1	US-08-290-919-11
4	44	41.5	20	3	US-08-578-674-22
5	44	41.5	20	4	US-09-498-346-22
6	44	41.5	66	3	US-08-578-674-5
7	44	41.5	66	4	US-09-498-346-5
8	43.5	41.0	1589	3	US-08-755-587-189
9	43	40.6	20	3	US-08-578-674-21
10	43	40.6	20	4	US-09-498-346-21
11	43	40.6	37	3	US-08-578-674-6
12	43	40.6	37	4	US-09-498-346-6
13	43	40.6	56	3	US-08-578-674-16
14	43	40.6	56	4	US-09-498-346-16
15	43	40.6	66	3	US-08-578-674-3
16	43	40.6	66	3	US-08-578-674-4
17	43	40.6	66	4	US-09-498-346-3
18	43	40.6	66	4	US-09-498-346-4
19	43	40.6	84	3	US-08-578-674-2
20	43	40.6	84	4	US-09-498-346-2
21	42	39.6	104	1	US-08-488-113B-153
22	42	39.6	104	1	US-08-477-484B-153
23	42	39.6	104	1	US-08-107-669D-17
24	42	39.6	104	1	US-08-472-788A-17
25	42	39.6	104	2	US-08-477-531B-17
26	42	39.6	104	2	US-08-646-360-153
27	42	39.6	104	2	US-08-082-842A-17

28	42	39.6	104	4	US-08-839-765-153	Sequence 153, App
29	42	39.6	104	4	US-09-136-389-153	Sequence 153, App
30	42	39.6	110	4	US-09-025-769B-33	Sequence 33, Appl
31	42	39.6	110	4	US-09-025-769B-53	Sequence 53, Appl
32	42	39.6	111	2	US-08-958-201-12	Sequence 12, Appl
33	42	39.6	111	2	US-08-958-201-14	Sequence 14, Appl
34	42	39.6	112	4	US-09-025-769B-19	Sequence 19, Appl
35	42	39.6	235	4	US-09-049-672A-10	Sequence 10, Appl
36	42	39.6	236	3	US-08-487-550-10	Sequence 10, Appl
37	42	39.6	1193	2	US-08-400-159-10	Sequence 10, Appl
38	42	39.6	1193	3	US-08-611-729A-10	Sequence 10, Appl
39	41	38.7	96	1	US-08-488-113B-158	Sequence 158, App
40	41	38.7	96	1	US-08-477-484B-158	Sequence 158, App
41	41	38.7	96	1	US-08-107-669D-22	Sequence 22, Appl
42	41	38.7	96	1	US-08-472-788A-22	Sequence 22, Appl
43	41	38.7	96	2	US-08-477-531B-22	Sequence 22, Appl
44	41	38.7	96	2	US-08-646-360-158	Sequence 158, App
45	41	38.7	96	2	US-08-082-842A-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-08-290-919-3
; Sequence 3, Application US/08290919
; Patent No. 5720959
; GENERAL INFORMATION:
; APPLICANT: HOLDER, ANTHONY A.
; APPLICANT: BLACKMAN, MICHAEL J.
; APPLICANT: CHAPPEL, JONATHAN A.
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
; TITLE OF INVENTION: VACCINE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,919
; FILING DATE: 04-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9203821.5
; FILING DATE: 22-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/00367
; FILING DATE: 22-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 212242/HCM/MJL/6BC8/
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 53 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site

```

; LOCATION: 1
; OTHER INFORMATION: /label= X
; OTHER INFORMATION: /note= "X" = M and N, or N"
; US-08-290-919-3

Query Match 100.0%; Score 106; DB 1; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 19; Conservative : 0; Mismatches 0; Indels

QY 1 EDGSGNGKKITCECTKPD 19
    |||||
DB 22 EDGSGNGKKITCECTKPD 40
    |||||

RESULT 2
US-08-290-919-4
; Sequence 4, Application US/08290919
; Patent No. 5720959
; GENERAL INFORMATION:
; APPLICANT: HOLDER, ANTHONY A.
; APPLICANT: BLACKMAN, MICHAEL J.
; APPLICANT: CHAPPEL, JONATHAN A.
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
; TITLE OF INVENTION: VACCINE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,919
; FILING DATE: 04-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9203821.5
; FILING DATE: 22-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/00367
; FILING DATE: 22-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 212242/HCM/MJL/5BC8/
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 53 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /label= X
; OTHER INFORMATION: /note= "X" = M and N, or N"
; US-08-290-919-4

```

Query Match 87.7%; Score 93; DB 1; Length 53;
Best Local Similarity 89.5%; Pred. No. 9.1e-07;

Matches 17: Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EDGSGNGKKITCECTKPDs 19
||||| |||||||||

Db 22 EDGSSRRKKITCECTKPDs 40

RESULT 3

US-08-290-919-11

; Sequence 11, Application US/08290919

; Patent No. 5720959

; GENERAL INFORMATION:

; APPLICANT: HOLDER, ANTHONY A.

; APPLICANT: BLACKMAN, MICHAEL J.

; APPLICANT: CHAPPEL, JONATHAN A.

; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA

; TITLE OF INVENTION: VACCINE

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CUSHMAN DAREY & CUSHMAN, L.L.P.

; STREET: 1100 NEW YORK AVENUE, N.W.

; CITY: WASHINGTON

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20005-3918

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/290,919

; FILING DATE: 04-OCT-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9203821.5

; FILING DATE: 22-FEB-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/GB93/00367

; FILING DATE: 22-FEB-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: KOKULIS, PAUL N.

; REGISTRATION NUMBER: 16,773

; REFERENCE/DOCKET NUMBER: 212242/HCM/MJL/6BC8/

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 861-3000

; TELEFAX: (202) 822-0944

; TELEX: 6714627 CUSH

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 106 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-290-919-11

```

Query Match      44.3%;   Score 47;   DB 1;   Length 106;
Best Local Similarity 50.0%;   Pred. No. 6;
Matches      8;   Conservative      2;   Mismatches      6;   Indels      0;   Gaps      0;

QY      2   DSGSNGKKIITCECTKP 17
      :| | | | | | | |
Db      76 _ESTENSKKIICTCKREP 91

RESULT      4
US-08-578-674-22
; Sequence 22, Application US/08578674
; Patent No. 6077827
; GENERAL INFORMATION:
; APPLICANT: KOLBE, Hanno V.J.

```

; APPLICANT: RASMUSSEN, Ulla B.
; APPLICANT: KREIL, Gunther
; APPLICANT: ACHSTETTER, Tilman
; TITLE OF INVENTION: FAMILY OF PEPTIDES KNOWN AS XENOXINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/578.674

; FILING DATE: 28-DEC-1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 90 07901

; FILING DATE: 29-JUN-1993

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94 00202

; FILING DATE: 11-JAN-1994

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9400062.9

; FILING DATE: 11-JAN-1994

; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M.

; REGISTRATION NUMBER: 36.607

; REFERENCE/DOCKET NUMBER: 017753-071

; TELEPHONE: (703) 836-6620

; TELEFAX: (703) 836-2021

; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20

; OTHER INFORMATION: /note= "analysed portion of
; OTHER INFORMATION: xenoxin-3, alkylated (3)"

US-08-578-674-22

Query Match 41.5%; Score 44; DB 3; Length 20;
Best Local Similarity 53.3%; Pred. NO. 2.9;
Matches 8; Conservative 3; Mismatches 4; Indels 0;
Gaps 0;

QY 5 SNGKKITCCTKPS 19
Db 6 ANGVMTOECAKEDT 20

RESULT 5

US-09-498-346-22

; Sequence 22, Application US/09498346

; Patent No. 6277822

; GENERAL INFORMATION:
; APPLICANT: KOLBE, Hanno V.J.

; APPLICANT: RASMUSSEN, Ulla B.

; APPLICANT: KREIL, Gunther

; APPLICANT: ACHSTETTER, Tilman

; TITLE OF INVENTION: FAMILY OF PEPTIDES KNOWN AS XENOXINS

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria

; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498.346

; FILING DATE: 04-FEB-2000

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/578,674

; FILING DATE: 28-DEC-1995

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 90 07901

; FILING DATE: 29-JUN-1993

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94 00202

; FILING DATE: 11-JAN-1994

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9400062.9

; FILING DATE: 11-JAN-1994

; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M.

; REGISTRATION NUMBER: 36.607

; REFERENCE/DOCKET NUMBER: 017753-071

; TELEPHONE: (703) 836-6620

; TELEFAX: (703) 836-2021

; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20

; OTHER INFORMATION: /note= "analysed portion of
; OTHER INFORMATION: xenoxin-3, alkylated (3)"

US-09-498-346-22

Query Match 41.5%; Score 44; DB 4; Length 20;
Best Local Similarity 53.3%; Pred. No. 2.9;
Matches 8; Conservative 3; Mismatches 4; Indels 0;
Gaps 0;

QY 5 SNGKKITCCTKPS 19
Db 6 ANGVMTOECAKEDT 20

RESULT 6

US-08-578-674-5

; Sequence 5, Application US/08578674

; Patent No. 6077827

; GENERAL INFORMATION:
; APPLICANT: KOLBE, Hanno V.J.

; APPLICANT: RASMUSSEN, Ulla B.

; APPLICANT: KREIL, Gunther

; APPLICANT: ACHSTETTER, Tilman

; TITLE OF INVENTION: FAMILY OF PEPTIDES KNOWN AS XENOXINS

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria

STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/578,674
FILING DATE: 28-DEC-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90 07901
FILING DATE: 29-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94 00202
FILING DATE: 11-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9400062.9
FILING DATE: 11-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meuth, Donna M.
REGISTRATION NUMBER: 36,607
REFERENCE/DOCKET NUMBER: 017753-071
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 66 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-578-674-5

Query Match 41.5%; Score 44; DB 3; Length 66;
Best Local Similarity 53.3%; Pred. No. 9.9;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 SNGKKITCECTKPD 19
:|||:|:|:|:|:
Db 8 ANGVKMTQECAKEDT 22

RESULT 7
US-09-498-346-5
Sequence 5, Application US/09498346
Patent No. 6277822
GENERAL INFORMATION:
APPLICANT: KOLBE, Hanno V.J.
APPLICANT: RASMUSSEN, Ulla B.
APPLICANT: KRELL, Gunther
APPLICANT: ACHSTER, Tilman
TITLE OF INVENTION: FAMILY OF PEPTIDES KNOWN AS XENOXINS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,346
FILING DATE: 04-FEB-2000

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/578,674
FILING DATE: 28-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90.07901
FILING DATE: 29-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94 00202
FILING DATE: 11-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9400062.9
FILING DATE: 11-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meuth, Donna M.
REGISTRATION NUMBER: 36,607
REFERENCE/DOCKET NUMBER: 017753-071
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 66 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-498-346-5

Query Match 41.5%; Score 44; DB 4; Length 66;
Best Local Similarity 53.3%; Pred. No. 9.9;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 SNGKKITCECTKPD 19
:|||:|:|:|:|:
Db 8 ANGVKMTQECAKEDT 22

RESULT 8
US-08-755-587-189
Sequence 189, Application US/08755587
Patent No. 6045997
GENERAL INFORMATION:
APPLICANT: Futreal, Phillip A
APPLICANT: Wooster, Richard F
APPLICANT: Ashworth, Alan
APPLICANT: Stratton, Michael R
TITLE OF INVENTION: Materials and methods relating to the
TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer
NUMBER OF SEQUENCES: 222
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson
STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
CITY: Raleigh
STATE: NC
COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/755,587
FILING DATE: 25-NOV-1996
PRIOR APPLICATION DATA: GB 9523959.6
APPLICATION NUMBER: GB 9523959.6
FILING DATE: 23-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9525555.0
FILING DATE: 14-DEC-1995
PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9617961.9
FILING DATE: 28-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kenneth D Sibley
REGISTRATION NUMBER: 31.665
REFERENCE/DOCKET NUMBER: 5405-135
INFORMATION FOR SEQ ID NO: 189:
SEQUENCE CHARACTERISTICS:
LENGTH: 1589 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-755-587-189

Query Match 41.0%; Score 43.5; DB 3; Length 1589;
Best Local Similarity 50.0%; Pred. No. 3e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 2 DSGSN-GKKITCECKRPD 18
Db 234 DSGNNFAFOVTKNCKRPD 251

RESULT 9
US-08-578-674-21
Sequence 21, Application US/08578674
Patent No. 6077827
GENERAL INFORMATION:
APPLICANT: KOLBE, Hanno V.J.
APPLICANT: RASMUSSEN, Ulla B.
APPLICANT: KREIL, Gunther
APPLICANT: ACHSTETTER, Tilman
TITLE OF INVENTION: FAMILY OF PEPTIDES KNOWN AS XENOXINS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/578,674
FILING DATE: 28-DEC-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90 07901
FILING DATE: 29-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94 00202
FILING DATE: 11-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9400062.9
FILING DATE: 11-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meuth, Donna M.
REGISTRATION NUMBER: 36.607
REFERENCE/DOCKET NUMBER: 017753-071
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

FEATURE:
NAME/KEY: Peptide
LOCATION: 1..20
OTHER INFORMATION: /note= "analysed portion of
OTHER INFORMATION: xenoxin-3, alkylated (2)"
US-08-578-674-21
Query Match 40.6%; Score 43; DB 3; Length 20;
Best Local Similarity 57.1%; Pred. No. 4;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 5 SNGKKITCECKRPD 18
Db 7 ANGKMTQCEAKED 20
RESULT 10
US-09-498-346-21
Sequence 21, Application US/09498346
Patent No. 6277822
GENERAL INFORMATION:
APPLICANT: KOLBE, Hanno V.J.
APPLICANT: RASMUSSEN, Ulla B.
APPLICANT: KREIL, Gunther
APPLICANT: ACHSTETTER, Tilman
TITLE OF INVENTION: FAMILY OF PEPTIDES KNOWN AS XENOXINS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,346
FILING DATE: 04-FEB-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/578,674
FILING DATE: 28-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90 07901
FILING DATE: 29-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94 00202
FILING DATE: 11-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9400062.9
FILING DATE: 11-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meuth, Donna M.
REGISTRATION NUMBER: 36.607
REFERENCE/DOCKET NUMBER: 017753-071
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..20

OTHER INFORMATION: /note= "analysed portion of
OTHER INFORMATION: xenoxin-3, alkylated (2)"
US-09-498-346-21

Query Match 40.6%; Score 43; DB 4; Length 20;
Best Local Similarity 57.1%; Pred. No. 4;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 5 SNGKKTCECTKPD 18
Db 7 ANGKMTQECACKED 20

RESULT 11
US-08-578-674-6
; Sequence 6, Application US/08578674
; Patent No. 6077827
; GENERAL INFORMATION:
; APPLICANT: KOLBE, Hanno V.J.
; APPLICANT: RASMUSSEN, Ulla B.
; APPLICANT: KREIL, Gunther
; APPLICANT: ACHSTETTER, Tilman
; TITLE OF INVENTION: FAMILY OF PEPTIDES KNOWN AS XENOXINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/578,674
; FILING DATE: 28-DEC-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 90 07901
; FILING DATE: 29-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94 00202
; FILING DATE: 11-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9400062.9
; FILING DATE: 11-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M.
; REGISTRATION NUMBER: 36,607
; REFERENCE/DOCKET NUMBER: 017753-071
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..37
; OTHER INFORMATION: /note= "analysed portion of
; OTHER INFORMATION: xenoxine-1, alkylated"
US-08-578-674-6

Query Match 40.6%; Score 43; DB 3; Length 37;

Best Local Similarity 53.3%; Pred. No. 7.6;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 5 SNGKKTCECTKPD 19
Db 8 ANGKMTQECACKED 22

RESULT 12
US-09-498-346-6
; Sequence 6, Application US/09498346
; Patent No. 6277822
; GENERAL INFORMATION:
; APPLICANT: KOLBE, Hanno V.J.
; APPLICANT: RASMUSSEN, Ulla B.
; APPLICANT: KREIL, Gunther
; APPLICANT: ACHSTETTER, Tilman
; TITLE OF INVENTION: FAMILY OF PEPTIDES KNOWN AS XENOXINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,346
; FILING DATE: 04-FEB-2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/578,674
; FILING DATE: 28-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 90 07901
; FILING DATE: 29-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94 00202
; FILING DATE: 11-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9400062.9
; FILING DATE: 11-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M.
; REGISTRATION NUMBER: 36,607
; REFERENCE/DOCKET NUMBER: 017753-071
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..37
; OTHER INFORMATION: /note= "analysed portion of
; OTHER INFORMATION: xenoxine-1, alkylated"
US-09-498-346-6

Query Match 40.6%; Score 43; DB 4; Length 37;
Best Local Similarity 53.3%; Pred. No. 7.6;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 SNGKKITCECTKPD 19
:|||:| | | | |
Db 8 ANGKMTQECAREDT 22

RESULT 13

US-08-578-674-16
; Sequence 16, Application US/08578674
; Patent No. 6077827

GENERAL INFORMATION:

APPLICANT: KOLBE, Hanno V.J.
APPLICANT: RASMUSSEN, Ulla B.
APPLICANT: KREIL, Guntherin
APPLICANT: ACHSTETTER, Tilman
TITLE OF INVENTION: FAMILY OF PEPTIDES KNOWN AS XENOXINS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/578,674
FILING DATE: 28-DEC-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90 07901
FILING DATE: 29-JUN-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94 00202

PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9400062.9

FILING DATE: 11-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meuth, Donna M.

REGISTRATION NUMBER: 36,607

REFERENCE/DOCKET NUMBER: 017753-071

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 56 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Peptide

LOCATION: 1..56

OTHER INFORMATION: /note= "analysed portion of

OTHER INFORMATION: xenoxin-2, alkylated"

US-08-578-674-16

Query Match 40.6%; Score 43; DB 3; Length 56;
Best Local Similarity 53.3%; Pred. No. 12;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 SNGKKITCECTKPD 19
:|||:| | | | |

Db

RESULT 14

US-08-578-674-3
; Sequence 3, Application US/08578674
; Patent No. 6077827

US-09-498-346-16
; Sequence 16, Application US/09498346
; Patent No. 6277822

GENERAL INFORMATION:

APPLICANT: KOLBE, Hanno V.J.
APPLICANT: RASMUSSEN, Ulla B.
APPLICANT: KREIL, Guntherin
APPLICANT: ACHSTETTER, Tilman
TITLE OF INVENTION: FAMILY OF PEPTIDES KNOWN AS XENOXINS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,346
FILING DATE: 04-FEB-2000

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/578,674
FILING DATE: 28-DEC-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90 07901

PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94 00202

FILING DATE: 11-JAN-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9400062.9

FILING DATE: 11-JAN-1994

ATTORNEY/AGENT INFORMATION:
NAME: Meuth, Donna M.

REGISTRATION NUMBER: 36,607

REFERENCE/DOCKET NUMBER: 017753-071

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 56 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Peptide

LOCATION: 1..56

OTHER INFORMATION: /note= "analysed portion of

OTHER INFORMATION: xenoxin-2, alkylated"

US-09-498-346-16

Query Match 40.6%; Score 43; DB 4; Length 56;
Best Local Similarity 53.3%; Pred. No. 12;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 SNGKKITCECTKPD 19
:|||:| | | | |

Db

Search completed: January 29, 2002, 10:24:05
Job time: 511 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:26:40 ; Search time 144.96 Seconds
(without alignments)
9.984 Million cell updates/sec

Title: US-09-763-397A-12

Perfect score: 106

Sequence: 1 EDGSGNGKRTCECTKPD 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106	100.0	400	2 A45545	major merozoite su
2	106	100.0	651	2 S47282	merozoite surface
3	106	100.0	1631	1 SA20K1	major merozoite su
4	106	100.0	1639	2 S05603	major merozoite su
5	106	100.0	1726	1 SA20GM	major merozoite su
6	106	100.0	1726	2 A45948	major merozoite su
7	102	96.2	1640	2 A24594	probable major sur
8	93	87.7	1701	2 A54498	major merozoite su
9	93	87.7	1701	2 A26868	major merozoite su
10	51	48.1	995	2 S03558	hypothetical prote
11	47	44.3	680	2 A28121	major merozoite su
12	47	44.3	1772	2 A45532	major merozoite su
13	44	41.5	167	2 T45261	hypothetical prote
14	44	41.5	341	2 S55627	hypothetical prote
15	44	41.5	468	2 B75053	hypothetical prote
16	44	41.5	2301	1 GNNVTM	genome polypeptid
17	44	41.5	2303	1 GNNVTM	genome polypeptid
18	44	41.5	2303	1 GNNVTM	genome polypeptid
19	44	41.5	2303	1 GNNVTM	genome polypeptid
20	44	41.5	3839	2 T48799	related to TOM1 pr
21	43.5	41.0	3328	2 T30835	breast cancer tumo
22	43.5	41.0	3329	2 T42205	breast cancer susc
23	43.5	41.0	3329	2 T30904	breast cancer tumo
24	43	40.6	84	2 T51698	pre-xenoxin-1 - Af
25	43	40.6	609	2 S45930	probable amino aci
26	42.5	40.1	171	1 CSRP	peptidylprolyl iso
27	42.5	40.1	263	2 G84547	hypothetical prote
28	42.5	40.1	2643	2 T29149	hypothetical prote
29	42	39.6	99	2 S36055	Ig lambda chain -

30	42	39.6	99	2 S36056	Ig lambda chain -
31	42	39.6	110	2 S51149	antibody light cha
32	42	39.6	110	2 S57412	Ig lambda chain V-
33	42	39.6	111	2 S36281	Ig lambda chain V
34	42	39.6	111	2 S38497	Ig lambda chain -
35	42	39.6	111	2 S46396	Ig lambda chain V
36	42	39.6	111	2 S46397	Ig lambda chain V
37	42	39.6	112	2 S31515	Ig lambda chain V
38	42	39.6	112	2 S44105	Ig lambda chain V-
39	42	39.6	112	2 S44123	Ig lambda chain V-
40	42	39.6	112	2 S46395	Ig lambda chain V
41	42	39.6	118	2 S04519	Ig lambda chain pr
42	42	39.6	143	2 T12144	hypothetical prote
43	42	39.6	230	2 S26199	plastoquinol--plas
44	42	39.6	235	2 S25758	Ig lambda chain -
45	42	39.6	253	2 T00967	hypothetical prote

ALIGNMENTS

RESULT 1
A45545 major merozoite surface antigen - malaria parasite (Plasmodium falciparum) (fragments
C:Species: Plasmodium falciparum
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jun-2000
C:Accession: A45545
R:Blackman, M.J.; Ling, I.T.; Nicholas, S.C.; Holder, A.A.
Mol. Biochem. Parasitol. 49, 29-33, 1991
A:Title: Proteolytic processing of the Plasmodium falciparum merozoite surface proteol
A:Reference number: A45545; MUID:92131048
A:Accession: A45545
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-400 <BLA>
A:Note: sequence extracted from NCBI backbone (NCBIN:77612, NCBI:77621)
C:Superfamily: major merozoite surface antigen
C:Keywords: glycoprotein; merozoite; surface antigen; tandem repeat

Query Match 100.0%; Score 106; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 5.9e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKRTCECTKPD 19
DB 351 EDGSGNGKRTCECTKPD 369

RESULT 2
S47282 merozoite surface antigen 1 - malaria parasite (Plasmodium falciparum) (strain RO-71)
C:Species: Plasmodium falciparum
A:Variety: strain RO-71
C:Date: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 09-Jun-2000
C:Accession: S47282
R:Tolle, R.; Bujard, H.; Cooper, J.A.
submitted to the EMBL Data Library, July 1994
A:Description: Plasmodium falciparum: recombination within the C-terminal region of m
A:Reference number: S47282
A:Accession: S47282
A:Molecule type: DNA
A:Residues: 1-651 <TOL>
A:Cross-References: EMBL:235329; NID:9535257; PIDN:CAAB4558.1; PID:9535258
A:Experimental source: strain RO-71
C:Superfamily: major merozoite surface antigen
C:Keywords: glycoprotein; merozoite; surface antigen

Query Match 100.0%; Score 106; DB 2; Length 651;
Best Local Similarity 100.0%; Pred. No. 8.7e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGSGNGKKTCTCKPDS 19
|||||
Db 602 EDGSGNGKKTCTCKPDS 620

RESULT 3
SAZOGM
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (st
C:Species: Plasmodium falciparum
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 09-Jun-2000
C:Accession: A25120
R:McKAY, M.; Bone, N.; Hyde, J.E.; Scaife, J.; Certa, U.; Stunnenberg, H.; H
EMBO J. 4, 3823-3829, 1985
A:Title: Polymorphism of the precursor for the major surface antigens of Plasmodium falc
A:Reference number: A91030; MUID:86136024
A:Accession: A25120
A:Molecule type: DNA
A:Residues: 1-1631 <MAC>
C:Comment: The merozoite stages of different strains have strain-specific surface antigen
C:Keywords: P. falciparum has three stages: sporozoite, merozoite, and gametocyte. The me
C:Keywords: glycoprotein; merozoite; surface antigen; tandem repeat; transmembrane prote
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1631/Product: major merozoite surface antigen #status predicted <MAT>
F:67-84/Region: 3-residue repeats (S-G-T/P)
F:1614-1631/Domain: membrane anchor #status predicted <MBN>
F:97,259,755,759,835,911,955,1049,1156,1436,1563/Binding site: carbohydrate (Asn) (

Query Match 100.0%; Score 106; DB 1; Length 1631;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 EDGSGNGKKTCTCKPDS 19
|||||
Db 1582 EDGSGNGKKTCTCKPDS 1600

RESULT 4
S05603
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (st
N:Alternate names: gp195 surface antigen
C:Species: Plasmodium falciparum
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jun-2000
C:Accession: S05603; S04850
R:Myer, P.J.
submitted to the EMBL Data Library, April 1989
A:Reference number: S05603
A:Accession: S05603
A:Molecule type: mRNA
A:Residues: 1-1639 <MYL>
A:Cross-references: EMBL:X15063; NID:99896; PIDD:CA31363.1; PID:99897
R:Myer, P.J.
Nucleic Acids Res. 17, 5401, 1989
A:Title: Nucleotide and deduced amino acid sequence of the gp195 (MSA-1) gene from Plas
A:Reference number: S04850; MUID:89345116
A:Accession: S04850
A:Molecule type: mRNA
A:Residues: 1504-1639 <MYL2>
A:Cross-references: EMBL:X15063
C:Superfamily: major merozoite surface antigen
C:Keywords: glycoprotein; merozoite; surface antigen
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1639/Product: major merozoite surface antigen #status predicted <MAT>

Query Match 100.0%; Score 106; DB 2; Length 1639;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 EDGSGNGKKTCTCKPDS 19
|||||
Db 1590 EDGSGNGKKTCTCKPDS 1608

RESULT 5
SAZOGM
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)

N:Alternate names: 195K glycoprotein
C:Species: Plasmodium falciparum
C:Date: 30-Sep-1987 #sequence_revision 31-Mar-1991 #text_change 09-Jun-2000
C:Accession: A23386; S06361
R:Weber, J.L.; Leininger, W.M.; Lyon, J.A.
Nucleic Acids Res. 14, 3311-3323, 1986
A:Title: Variation in the gene encoding a major merozoite surface antigen of the huma
A:Reference number: A23386; MUID:86205236
A:Accession: A23386
A:Molecule type: DNA
A:Residues: 1-1104 <WEB1>
A:Cross-references: EMBL:X03831
R:Weber, J.L.; Sim, B.K.L.; Lyon, J.A.; Wolff, R.
Nucleic Acids Res. 16, 1206, 1988
A:Title: Merozoite surface protein sequence from the Camp strain of the human malaria
A:Reference number: S06361; MUID:88143999
A:Accession: S06361
A:Molecule type: DNA
A:Residues: 1104-1726 <WEB2>
A:Cross-references: EMBL:X03831
C:Comment: The merozoite stages of different strains have strain-specific surface ant
C:Keywords: P. falciparum has three stages: sporozoite, merozoite, and gametocyte. The
C:Superfamily: major merozoite surface antigen
C:Keywords: glycoprotein; merozoite; surface antigen; tandem repeat
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1726/Product: major merozoite surface antigen #status predicted <MAT>
F:67-87,91-96,100-105,109-120/Region: 3-residue repeats (S-G-T)
F:757-765/Region: 3-residue repeats (T-E-E)
F:133,272,501,567,638,827,839,924,944,990,1016,1114,1221,1613,1658/Binding site: carb

Query Match 100.0%; Score 106; DB 1; Length 1726;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGSGNGKKTCTCKPDS 19
|||||
Db 1677 EDGSGNGKKTCTCKPDS 1695

RESULT 6
A45948
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jun-2000
C:Accession: A45948
R:Chang, S.P.; Kramer, K.J.; Yamaga, K.M.; Kato, A.; Case, S.E.; Siddiqui, W.A.
Exp. Parasitol. 67, 1-11, 1988
A:Title: Plasmodium falciparum: gene structure and hydropathy profile of the major me
A:Reference number: A45948; MUID:89005525
A:Accession: A45948
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1726 <CHA>
A:Cross-references: GB:M37213
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen

Query Match 100.0%; Score 106; DB 2; Length 1726;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 EDGSGNGKKTCTCKPDS 19
|||||
Db 1677 EDGSGNGKKTCTCKPDS 1695

RESULT 7
A:Accession: A24594
probable major surface antigen (83K, 19K, 42K) precursor - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 29-Aug-1987 #sequence_revision 29-Aug-1987 #text_change 09-Jun-2000
C:Accession: A24594
R:Holder: A.A.; Lockyer, M.J.; Odink, K.G.; Sandhu, J.S.; Riveros-Moreno, V.; Nicholls, Nature 317, 270-273, 1985
A:Title: Primary structure of the precursor to the three major surface antigens of Plasmodium falciparum
A:Reference number: A24594; MUID:86014355
A:Accession: A24594
A:Molecule type: DNA
A:Residues: 1-1640 <HOL>
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen

Query Match 96.2%; Score 102; DB 2; Length 1640;
Best Local Similarity 100.0%; Pred. No. 6,7e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDGSGNGKRTCTCKRPD 18
|||||
Db 1591 EDGSGNGKRTCTCKRPD 1608

RESULT 8
A:Accession: A54498
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (is C:Species: Plasmodium falciparum
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C:Accession: A54498
R:Peterson, M.G.; Coppel, R.L.; McIntyre, P.; Langford, C.J.; Woodrow, G.; Brown, G.V.; Mol. Biochem. Parasitol. 27, 291-302, 1988
A:Title: Variation in the precursor to the major merozoite surface antigens of Plasmodium falciparum
A:Reference number: A54498; MUID:88142999
A:Accession: A54498
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1701 <PRT>
A:Cross-references: GB:M19143; NID:q160412; PIDN:AAA29653.1; PID:q160413
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen

Query Match 87.7%; Score 93; DB 2; Length 1701;
Best Local Similarity 89.5%; Pred. No. 1.4e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGSGNGKRTCTCKRPPS 19
|||||
Db 1652 EDGSGNGKRTCTCKRPPS 1670

RESULT 9
A:Accession: A26868
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (st C:Species: Plasmodium falciparum
C:Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 09-Jun-2000
C:Accession: A26868
R:Tanabe, K.; Mackay, M.; Goman, M.; Scalfe, J.G. J. Mol. Biol. 195, 273-287, 1987
A:Title: Allelic dimorphism in a surface antigen gene of the malaria parasite Plasmodium falciparum
A:Reference number: A26868; MUID:88011243
A:Accession: A26868
A:Molecule type: DNA
A:Residues: 1-1701 <TRAN>
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1701/Product: major merozoite surface antigen #status predicted <MAN>

Query Match 87.7%; Score 93; DB 2; Length 1701;
Best Local Similarity 89.5%; Pred. No. 1.4e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDGSGNGKRTCTCKRPPS 19
|||||
Db 1652 EDGSGNGKRTCTCKRPPS 1670

RESULT 10
A:Accession: S50358
hypothetical protein YIL169c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein Y19402.07c
C:Species: Saccharomyces cerevisiae
C:Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 05-Nov-1999
C:Accession: S50358
R:Lyle, G.; Bowman, S.; Churcher, C. submitted to the EMBL Data Library, December 1994
A:Reference number: S50349
A:Accession: S50358
A:Molecule type: DNA
A:Residues: 1-995 <LYE>
A:Cross-references: GB:Z47047; EMBL:Z46921; NID:q603997; PID:q604006; GSPDB:GND0009;
C:Genetics:
A:Gene: MIPS:YIL169c
A:Map position: 9L

Query Match 48.1%; Score 51; DB 2; Length 995;
Best Local Similarity 47.4%; Pred. No. 11;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 1 EDGSGNGKRTCTCKRPPS 19
:|:|:|:|:|:|:|:|:|
Db 735 DDNGCTKTVTSECKETS 753

RESULT 11
A:Accession: A28121
major merozoite surface antigen - Plasmodium yoelii (fragment)
C:Species: Plasmodium yoelii
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 26-Aug-1999
C:Accession: A28121
R:Burns Jr., J.M.; Daly, T.M.; Valdivia, A.B.; Long, C.A. Proc. Natl. Acad. Sci. U.S.A. 85, 602-606, 1988
A:Title: The 3' portion of the gene for a Plasmodium yoelii merozoite surface antigen
A:Reference number: A28121; MUID:88124889
A:Accession: A28121
A:Molecule type: DNA
A:Residues: 1-680 <BUR>
A:Cross-references: GB:J03612; NID:q160678; PID:q160679
A:Experimental source: strain 17XL
A>Note: the authors translated the codon GTA for residue 429 as Leu
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen

Query Match 44.3%; Score 47; DB 2; Length 680;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 2 DSGSGNGKRTCTCKRP 17
:|:|:|:|:|:|:|:|:|
Db 632 ESTENSKRTCTCKRP 647

RESULT 12
A:Accession: A45532
major merozoite surface antigen precursor - Plasmodium yoelii
C:Species: Plasmodium yoelii
C:Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 26-Aug-1999
C:Accession: A45532; A45531
R:Lewis, A.P.

Mol. Biochem. Parasitol. 36, 271-282, 1989
A:Title: Cloning and analysis of the gene encoding the 230-kilodalton merozoite surface
A:Reference number: A45532; MUID:90014981
A:Accession: A45532
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1772 <LEM>
A:Cross-references: GB:J04668; NID:g160492; PID:g160493
R:Daly, T.M.; Burns Jr., J.M.; Long, C.A.
Mol. Biochem. Parasitol. 36, 283-285, 1989
A:Title: Precursor to the major merozoite surface antigen of Plasmodium yoelii: cloning
A:Reference number: A45531; MUID:90014982
A:Accession: A45531
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 454-1094 <DAL>
A:Cross-references: GB:J03975; NID:g160081; PID:g160082
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen

Query Match 44.3%; Score 47; DB 2; Length 1772;
Best Local Similarity 50.0%; Pred. No. 67;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
OY 2 DSGSNGKKTCTCKP 17
: | | | | | | | | | |
Db 1724 ESTENSKKIKCTCKP 1739

RESULT 13
T43261
hypothetical protein [imported] - Methanobacterium thermoautotrophicum (fragment)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C:Accession: T45261
R:Harms, U.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z22950
A:Accession: T45261
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-167 <HAR>
A:Cross-references: EMBL:X97589; PIDN:CA66200.1
A:Experimental source: strain Marburg; DSM 2133

Query Match 41.5%; Score 44; DB 2; Length 167;
Best Local Similarity 46.7%; Pred. No. 29;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
OY 3 SGSGNGKKTCTCKP 17
: | | | | | | | | | |
Db 7 AGGEGKRLGLECEKP 21

RESULT 14
S55627
hypothetical protein 33 - equine herpesvirus 2
C:Species: equine herpesvirus 2
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 26-Aug-1999
C:Accession: S55627
R:Telford, E.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.
J. Mol. Biol. 249, 520-528, 1995
A:Title: The DNA sequence of equine herpesvirus 2.
A:Reference number: S55594; MUID:95302501
A:Accession: S55627
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-341 <TEL>
A:Cross-references: GB:U20824; NID:g695172; PIDN:AC13820.1; PID:g695205
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995
C:Superfamily: Epstein-Barr virus BGLF2 protein

Query Match 41.5%; Score 44; DB 2; Length 341;
Best Local Similarity 57.1%; Pred. No. 51;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
OY 2 DSGSNGKKTCTCKP 15
: | | | | | | | | | |
Db 233 DPGSDGVSTCLCT 246

RESULT 15
B75053
dint related PAB1439 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: B75053
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome s
A:Reference number: A75001
A:Accession: B75053
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-468 <RAW>
A:Cross-references: GB:AJ248287; GB:AL096836; NID:g5458657; PIDN:CAB50319.1; PID:g545
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB1439
C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0709

Query Match 41.5%; Score 44; DB 2; Length 468;
Best Local Similarity 62.5%; Pred. No. 65;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
OY 3 SGSGNGKKTCTCKRP 18
: | | | | | | | | | |
Db 221 SGKAGLKTLEDLKP 236

Search completed: January 29, 2002, 10:26:41
Job time: 652 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:13:42 : Search time 80.65 seconds
(without alignments)
8.638 Million cell updates/sec

Title: US-09-763-397A-12

Perfect score: 106

Sequence: 1 EDGSGNGKITCTKTPDS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	106	100.0	1630	1 MSP1_PLAFK	P04932 plasmodium
2	106	100.0	1639	1 MSP1_PLAFW	P04933 plasmodium
3	106	100.0	1682	1 MSP1_PLAF3	P19598 plasmodium
4	106	100.0	1726	1 MSP1_PLAF3	P04934 plasmodium
5	106	100.0	1726	1 MSP1_PLAF3	P50495 plasmodium
6	93	87.7	1701	1 MSP1_PLAF3	P13819 plasmodium
7	93	87.7	1701	1 MSP1_PLAF3	P08569 plasmodium
8	51	48.1	995	1 Y109_YEAST	P40442 saccharomyc
9	47	44.3	1772	1 MSP1_PLAYO	P13828 plasmodium
10	45.5	42.9	795	1 CDB4_HUMAN	Q9V5E5 homo sapien
11	45	42.5	278	1 ATND_MOUSE	P97370 mus musculu
12	45	42.5	1343	1 TIM_DROVI	O17482 drosophila
13	44	41.5	66	1 XEN3_XENLA	P38952 xenopus lae
14	44	41.5	652	1 VE1_HPV70	P50765 human papil
15	44	41.5	2301	1 POLG_TMEVD	P13899 t genome po
16	44	41.5	2303	1 POLG_TMEVB	P08544 t genome po
17	44	41.5	2303	1 POLG_TMEVC	P08545 t genome po
18	43.5	41.0	3329	1 BRC2_MOUSE	P97929 mus musculu
19	43	40.6	66	1 XEN2_XENLA	P38951 xenopus lae
20	43	40.6	73	1 MT_DREPO	Q94550 dreissena p
21	43	40.6	84	1 XEN1_XENLA	Q09022 xenopus lae
22	43	40.6	609	1 BAP2_YEAST	P38084 saccharomyc
23	42.5	40.1	171	1 CYPH_BRANA	P24525 brassica na
24	42.5	40.1	950	1 CDAD_HUMAN	Q9Y510 homo sapien
25	42	39.6	230	1 UCR1_FEA	P26291 pisum sativ
26	42	39.6	527	1 RAG2_HUMAN	P55895 homo sapien
27	42	39.6	527	1 RAG2_MOUSE	P21784 mus musculu
28	41.5	39.2	419	1 VSY1_TRYCO	P20948 trypanosoma
29	41	38.7	66	1 MTCD_HELPO	P33187 helix pomat
30	41	38.7	111	1 LV2B_HUMAN	P01705 homo sapien
31	41	38.7	111	1 LV2B_HUMAN	P01707 homo sapien
32	41	38.7	111	1 LV2G_HUMAN	P01710 homo sapien
33	41	38.7	111	1 LV2H_HUMAN	P01711 homo sapien

RESULT 1

MSPI_PLAFK

ID

MSPI_PLAFK

STANDARD;

PRT: 1630 AA.

DT

13-AUG-1987 (Rel. 05, Created)

DT

01-FEB-1996 (Rel. 33, Last sequence update)

DT

01-OCT-1996 (Rel. 34, Last annotation update)

DE

MEEROZITE SURFACE PROTEIN 1 PRECURSOR (MEEROZITE SURFACE ANTIGENS)

DE

(PMMSA) (P190).

GN

MSP-1.

OS

Plasmodium falciparum (isolate K1 / Thailand).

OC

Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX

NCBI_TaxID=5839;

RN

[1]

RP

SEQUENCE FROM N.A.

RX

MEDLINE=86136024; PubMed=3004972;

RA

Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,

RA

Stunnenberg H., Bujard H.;

RT

*Polymorphism of the precursor for the major surface antigens of

RT

Plasmodium falciparum merozoites: studies at the genetic level.;

RL

EMBO J. 4:3823-3829(1985).

RN

[2]

RP

REVISIONS, SEQUENCE FROM N.A.

RA

Pan W., Tolle R., Bujard H.;

RL

Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.

CC

-!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR

CC

(POTENTIAL).

CC

-!- PTM: MEEROZITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42

CC

KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF

CC

MEEROZITES. THE MATURATION TAKE PLACE DURING SCHIZONT.

34 41 38.7 112 1 LV2K_HUMAN P04209 homo sapien
35 41 38.7 421 1 ECB2_HALEL O52250 halomonas e
36 41 38.7 454 1 MUC_HUMAN P01871 homo sapien
37 41 38.7 494 1 AMY2_SALTY P26613 salmonella
38 41 38.7 527 1 RAG2_RABIT P34089 oryctolagus
39 41 38.7 1328 1 F1NC_MOUSE O91289 pleurodeles
40 41 38.7 5376 1 ZAN_MOUSE O88799 mus musculus
41 40.5 38.2 544 1 RGSE_RAT O08773 rattus norv
42 40.5 38.2 661 1 F13B_HUMAN O9V512 homo sapien
43 40.5 38.2 948 1 CDAA_HUMAN P01703 homo sapien
44 40 37.7 103 1 LVIE_HUMAN
45 40 37.7 111 1 LV2F_HUMAN P01709 homo sapien

ALIGNMENTS

FT CARBOHYD 955 955 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1156 1156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1436 1436 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1517 1517 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1630 AA; 187289 MW; ADBDEC3CE0A46322 CRC64;

Query Match 100.0%; Score 106; DB 1; Length 1630;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1.EDSGSNGKKTCTCKPDS 19
|||||
Db 1581 EDGSNGKKTCTCKPDS 1599

RESULT 2

MSPL_PLAFW STANDARD; PRT; 1639 AA.
AC P04933;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
DE (PMMSA) (P195).
GN MSP-1.
OS Plasmodium falciparum (isolate Wellcome).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5848;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86014355; PubMed=2995820;
RA Holder A.A.; Lockyer M.J.; Odink K.G.; Sandhu J.S.; Riveros-Moreno V.,
RA Nicholls S.C.; Hillman Y.; Davey L.S.; Tizard M.L.V.; Schwarz R.T.,
RA Freeman R.R.;
RT "Primary structure of the precursor to the three major surface
antigens of Plasmodium falciparum merozoites.";
RL Nature 317:270-273(1985).
RN [2]
RP REVISIONS.
RA Holder A.A.;
RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
(POTENTIAL).
CC -!- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42
KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X02919; CAA26676.1; -
CC PIR; A24594; A24594.
CC InterPro; IPR000361; EGF-like.
CC Pfam; PF00008; EGF; 1.
CC Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
CC Transmembrane; GPI-anchor.
CC SIGNAL 1 19
CC CHAIN 1 19
CC CARBOHYD 116 116 MEROZOITE SURFACE PROTEIN 1.
CC CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 764 764 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 768 768 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 783 783 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 844 844 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 920 920 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 964 964 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1058 1058 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1174 1174 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1445 1445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1526 1526 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1639 AA; 187618 MW; 2C255B6616C87F6E CRC64;

Query Match 100.0%; Score 106; DB 1; Length 1639;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1.EDSGSNGKKTCTCKPDS 19
|||||
Db 1590 EDGSNGKKTCTCKPDS 1608

RESULT 3

MSPL_PLAF3 STANDARD; PRT; 1682 AA.
AC P19598; Q25921;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
DE (PMMSA) (P190).
GN MSP-1.
OS Plasmodium falciparum (isolate ro-33 / Ghana).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5834;
RN [1]
RP SEQUENCE OF 1-1061 FROM N.A.
RX MEDLINE=88166657; PubMed=3327688;
RA Certa U.; Rotmann D.; Matile H.; Reber-Liske R.;
RT "A naturally occurring gene encoding the major surface antigen
precursor p190 of Plasmodium falciparum lacks tripeptide repeats.";
RL EMBO J. 6:4137-4142(1987).
RN [2]
RP SEQUENCE OF 1032-1682 FROM N.A.
RX MEDLINE=95354793; PubMed=7628566;
RA Tolle R.; Bujard H.; Cooper J.A.;
RT "Plasmodium falciparum: variations within the C-terminal region of
merozoite surface antigen-1.";
RL Exp. Parasitol. 81:47-54(1995).
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
(POTENTIAL).
CC -!- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42
KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; M35727; AAA29715.1; -
CC EMBL; Y00087; CAA68280.1; -
CC EMBL; Z53326; CAA84555.1; -
CC PIR; S06286; S06286.
CC InterPro; IPR000561; EGF-like.
CC Pfam; PF00008; EGF; 1.
CC Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
CC Transmembrane; GPI-anchor.
CC SIGNAL 1 19
CC CHAIN 1 19
CC CARBOHYD 20 1682 MEROZOITE SURFACE PROTEIN 1.
CC CARBOHYD 1666 1682 MEMBRANE ANCHOR (POTENTIAL).
CC TRANSMEM 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 462 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 528 528 N-LINKED (GLCNAC. . .) (POTENTIAL).

```
FT CARBOHYD 599 785 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 881 785 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 901 881 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 947 901 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1071 947 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1178 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1569 1178 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1682 AA; 192462 MW; C82A1E159948CAD6 CRC64;

Query Match 100.0%; Score 106; DB 1; Length 1682;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDGSGNGKKITCECTKPD 19
Db 1633 EDGSGNGKKITCECTKPD 1651

RESULT 4
MSPL_PLAFC
ID MSPL_PLAFC STANDARD; PRT; 1726 AA.
AC P04934;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
DE (PMMSA) (P195).
GN MSP-1.
OS Plasmodium falciparum (isolate Camp / Malaysia).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5835;
RN [1]
RP SEQUENCE OF 1-1103 FROM N.A.
RX MEDLINE=86205236; PubMed=3517809;
RA Weber J.L., Leininger W.M., Lyon J.A.;
RT "Variation in the gene encoding a major merozoite surface antigen of the human malaria parasite Plasmodium falciparum.";
RL Nucleic Acids Res. 14:3311-3323(1986).
RN [2]
RP SEQUENCE OF 1104-1726 FROM N.A.
RX MEDLINE=88143999; PubMed=3278296;
RA Weber J.L., Sim B.K.L., Lyon J.A., Wolff R.;
RT "Merozoite surface protein sequence from the Camp strain of the human malaria parasite Plasmodium falciparum.";
RL Nucleic Acids Res. 16:1206-1206(1988).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (POTENTIAL).
CC -1- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03831; CAA27446.1; .
DR PIR; A23386; SAZOGM
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein; Transmembrane; GPI-anchor.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1726 MEROZOITE SURFACE PROTEIN 1.
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT Transmembrane; GPI-anchor.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1726 MEROZOITE SURFACE PROTEIN 1.
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT CARBOHYD 638 827 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 924 827 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 944 944 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 990 944 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1016 990 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1114 1016 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1221 1114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1613 1221 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1726 AA; 196197 MW; DD8AD45FA352BCF3 CRC64;

Query Match 100.0%; Score 106; DB 1; Length 1726;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDGSGNGKKITCECTKPD 19
Db 1677 EDGSGNGKKITCECTKPD 1695

RESULT 5
MSPL_PLAFC
ID MSPL_PLAFC STANDARD; PRT; 1726 AA.
AC P50495;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
DE (PMMSA) (P195).
GN MSP-1.
OS Plasmodium falciparum (isolate Palo Alto / Uganda).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57270;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89005525; PubMed=3049134;
RA Chang S.P., Kramer K.J., Yamaga K.M., Kato A., Case S.E., Siddiqui W.A.;
RT "Plasmodium falciparum: gene structure and hydropathy profile of the major merozoite surface antigen (gp195) of the Uganda-Palo Alto isolate.";
RL Exp. Parasitol. 67:1-11(1988).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (POTENTIAL).
CC -1- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M37213; AAA29611.1; .
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein; Transmembrane; GPI-anchor.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1726 MEROZOITE SURFACE PROTEIN 1.
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT Transmembrane; GPI-anchor.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1726 MEROZOITE SURFACE PROTEIN 1.
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 638 567 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 827 638 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 924 827 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 944 924 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 990 944 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

•

```
FT CARBOHYD 919 919 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 965 965 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 991 991 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1089 1089 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1196 1196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1588 1588 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1701 AA; 193768 MW; 3FC2EC59AF96EA98 CRC64;

Query Match 87.7%; Score 93; DB 1; Length 1701;
Best Local Similarity 89.5%; Pred. No. 1.3e-06;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPD 19
|||||: |||||
Db 1652 EDGSGSRKKITCECTKPD 1670

RESULT 8
Y109_YEAST
ID Y109_YEAST STANDARD; PRT; 995 AA.
AC P40442.
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL 99.7 KDA PROTEIN IN SDL1 5'REGION PRECURSOR.
GN Y1169C OR Y19402.07C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churche C.M., Connor R., Copest T., Dear S., Devlin K., Fraser A.,
RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 246921; CAAB7023.1; -.
DR SGD; S0001431; Y1169C.
DR InterPro; IPR000122; Chemotaxis_transducer.
DR InterPro; IPR000727; T_SNARE.
KW Hypothetical protein; Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 995 HYPOTHETICAL PROTEIN Y1169C.
FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 664 664 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 995 AA; 99735 MW; F63E287A03F137EC CRC64;

Query Match 48.1%; Score 51; DB 1; Length 995;
Best Local Similarity 47.4%; Pred. No. 2.5;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPD 19
|||||: |||||
Db 735 DONGCWTKTVTSECKETS 753
```

```
RESULT 9
MSPL_PLAYO
ID MSPL_PLAYO STANDARD; PRT; 1772 AA.
AC P13828.
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
DE (PMMSA) (230 KDA).
DE MSP-1.
GN Plasmodium berghei yoelii.
OS Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5862;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90205979; PubMed=2320061;
RA Lewis A.P.;
RT "Sequence analysis upstream of the gene encoding the precursor to the
RT major merozoite surface antigens of Plasmodium yoelii.";
RL Mol. Biochem. Parasitol. 39:285-288(1990).
RN [2]
RP SEQUENCE OF 1093-1772 FROM N.A.
RC STRAIN=17XL;
RX MEDLINE=88124889; PubMed=2448778;
RA Burns J.M. Jr., Daly T.M., Vaidya A.B., Long C.A.;
RT "The 3' portion of the gene for a Plasmodium yoelii merozoite surface
RT antigen encodes the epitope recognized by a protective monoclonal
RT antibody.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:602-606(1988).
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
CC (POTENTIAL).
CC -!- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42
CC KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
CC MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03612; AAA29762.1; -.
DR EMBL; J04668; AAA29702.1; -.
DR PIR; A28121; A28121.
DR PIR; A45532; A45532.
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchor.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 1772 MEROZOITE SURFACE PROTEIN 1.
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 646 646 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 829 829 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1090 1090 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1408 1408 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1446 1446 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1541 1541 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1629 1629 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1680 1680 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 1521 1521 L -> V (IN REF. 2).
SQ SEQUENCE 1772 AA; 197230 MW; 9A6291658E90F45D CRC64;

Query Match 44.3%; Score 47; DB 1; Length 1772;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 DSGSGNGKKITCECTK 17
|||||: |||||
Db 1724 ESTENSKKIICTEK 1739
```

ID	ATND_MOUSE	STANDARD;	PRT;	278 AA.
AC	P97370;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	SODIUM/POTASSIUM-TRANSPORTING ATPASE BETA-3 CHAIN (SODIUM/POTASSIUM-DEPENDENT ATPASE BETA-3 SUBUNIT) (ATPB-3).			
GN	ATPB3.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALE/C; TISSUE=Retina;			
RX	MEDLINE=97157071; PubMed=9003452;			
RA	Besirli C.G., Gong T.-W.L., Lomax M.I.;			
RT	"Novel beta 3 isoform of the Na,K-ATPase beta subunit from mouse retina."			
RL	Biochim. Biophys. Acta 1350:21-26(1997).			
CC	-!- FUNCTION: THIS IS THE NON-CATALYTIC COMPONENT OF THE ACTIVE ENZYME, WHICH CATALYZES THE HYDROLYSIS OF ATP COUPLED WITH THE EXCHANGE OF NA AND K IONS ACROSS THE PLASMA MEMBRANE. THE EXACT FUNCTION OF THIS GLYCOPROTEIN IS NOT KNOWN.			
CC	-!- SUBUNIT: COMPOSED OF THREE SUBUNITS: ALPHA (CATALYTIC), BETA AND GAMMA.			
CC	-!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.			
CC	-!- TISSUE SPECIFICITY: WIDELY EXPRESSED.			
CC	-!- MISCELLANEOUS: THE BETA SUBUNIT SEEMS TO BE ENCODED BY A MULTIGENE FAMILY. EACH DIFFERENT SUBUNIT MAY HAVE SPECIALIZED FUNCTIONS.			
CC	-!- SIMILARITY: BELONGS TO THE NA+/K+ AND H+ ATPASES BETA CHAIN FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U59761; AAC00019.1; -			
DR	MGD; MGI:107788; Atplb3.			
DR	InterPro; IPR000402; Na_K_beta.			
DR	Pfam; PF00287; Na_K-ATPase; 1.			
DR	PROSITE; PS00390; ATPASE_NA_K_BETA.1; 1.			
KW	Sodium/potassium transport; Transmembrane; Glycoprotein;			
KW	Multigene family; Signal-anchor.			
FT	DOMAIN 1 35 CYTOPLASMIC (POTENTIAL).			
FT	TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).			
FT	DOMAIN 57 278 EXTRACELLULAR (POTENTIAL).			
FT	DISULFID 128 144 BY SIMILARITY.			
FT	DISULFID 154 170 BY SIMILARITY.			
FT	DISULFID 191 249 BY SIMILARITY.			
FT	CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).			
SQ	SEQUENCE 278 AA; 31775 MW; 98A46FF6EDA85AAC CRC64;			
QY	Query Match 42.5%; Score 45; DB 1; Length 278;			
Db	Best Local Similarity 61.5%; Pred. No. 6, 5;			
	Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;			
QY	2 DSGSNGKKITCEC 14			
Db	237 DSGLNKKEVTVEC 249			
RESULT 12				
ATND_MOUSE	STANDARD;			
AC	P97370;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	SODIUM/POTASSIUM-TRANSPORTING ATPASE BETA-3 CHAIN (SODIUM/POTASSIUM-DEPENDENT ATPASE BETA-3 SUBUNIT) (ATPB-3).			
GN	ATPB3.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALE/C; TISSUE=Retina;			
RX	MEDLINE=97157071; PubMed=9003452;			
RA	Besirli C.G., Gong T.-W.L., Lomax M.I.;			
RT	"Novel beta 3 isoform of the Na,K-ATPase beta subunit from mouse retina."			
RL	Biochim. Biophys. Acta 1350:21-26(1997).			
CC	-!- FUNCTION: THIS IS THE NON-CATALYTIC COMPONENT OF THE ACTIVE ENZYME, WHICH CATALYZES THE HYDROLYSIS OF ATP COUPLED WITH THE EXCHANGE OF NA AND K IONS ACROSS THE PLASMA MEMBRANE. THE EXACT FUNCTION OF THIS GLYCOPROTEIN IS NOT KNOWN.			
CC	-!- SUBUNIT: COMPOSED OF THREE SUBUNITS: ALPHA (CATALYTIC), BETA AND GAMMA.			
CC	-!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.			
CC	-!- TISSUE SPECIFICITY: WIDELY EXPRESSED.			
CC	-!- MISCELLANEOUS: THE BETA SUBUNIT SEEMS TO BE ENCODED BY A MULTIGENE FAMILY. EACH DIFFERENT SUBUNIT MAY HAVE SPECIALIZED FUNCTIONS.			
CC	-!- SIMILARITY: BELONGS TO THE NA+/K+ AND H+ ATPASES BETA CHAIN FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AF152497; AAD43758.1; -			
DR	HSSP; P15116; INCU.			
DR	InterPro; IPR002126; Cadherin.			
DR	Pfam; PF00028; cadherin; 5.			
DR	PRINTS; PR00205; CADHERIN.			
DR	SMART; SM00112; CA: 5.</			

15-JUL-1999 (Rel. 38, Created)
15-JUL-1999 (Rel. 38, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
TIMELESS PROTEIN.
TIM.
OS Drosophila virilis (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7244;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=98033379; PubMed=9365248;
RA Myers M.P., Rothenfluh A., Chang M., Young M.W.;
RT "Comparison of chromosomal DNA composing timeless in Drosophila
melanogaster and D. virilis suggests a new conserved structure for the
TIMELESS protein.";
RL Nucleic Acids Res. 25:4710-4714(1997).
RN [2]
RN SEQUENCE OF 1-1194 AND 1196-1343 FROM N.A.
RX MEDLINE=98186271; PubMed=9504927;
RA Ousley A., Zafarullah K., Chen Y., Emerson M., Hickman L., Sehgal A.;
RT "Conserved regions of the timeless (tim) clock gene in Drosophila
analyzed through phylogenetic and functional studies.";
RL Genetics 148:815-825(1998).
CC -!- FUNCTION: REQUIRED FOR THE PRODUCTION OF CIRCADIAN RHYTHMS. THE
BIOLOGICAL CYCLE DEPENDS ON THE RHYTHMIC FORMATION AND NUCLEAR
LOCALIZATION OF THE TIM-PER COMPLEX. LIGHT INDUCES THE DEGRADATION
OF TIM, WHICH PROMOTES ELIMINATION OF PER. NUCLEAR ACTIVITY OF THE
HETERODIMER COORDINATIVELY REGULATES PER. NUCLEAR ACTIVITY OF THE
THROUGH A NEGATIVE FEEDBACK LOOP. BEHAVES AS A NEGATIVE ELEMENT
IN CIRCADIAN TRANSCRIPTIONAL LOOP. DOES NOT APPEAR TO BIND DNA,
SUGGESTING INDIRECT TRANSCRIPTIONAL INHIBITION (BY SIMILARITY).
CC -!- SUBUNIT: FORMS A HETERODIMER WITH PERIOD (PER); THE COMPLEX THEN
TRANSLOCATES INTO THE NUCLEUS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR AT SPECIFIC PERIODS OF THE DAY.
FIRST ACCUMULATES IN THE PERINUCLEAR REGION ABOUT ONE HOUR BEFORE
TRANSLOCATION INTO THE NUCLEUS. INTERACTION WITH PER IS REQUIRED
FOR NUCLEAR LOCALIZATION (BY SIMILARITY).
CC -!- PTM: PHOSPHORYLATED WITH A CIRCADIAN RHYTHMICITY (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; AF032403; AAB94891.1; -;
DR EMBL; AF032402; AAB94891.1; JOINED.
DR EMBL; AF038502; AAB94930.1; -;
DR EMBL; AF040096; AAB94942.1; -;
DR FlyBase; FBgn0022834; Dvir\Tim.
KW Biological rhythms; Nuclear protein; Repeat; Phosphorylation.
FT REPEAT 480 554 PAS-1.
FT REPEAT 691 890 PAS-2.
FT DOMAIN 238 270 NECESSARY FOR NORMAL CIRCADIAN RHYTHM (BY
SIMILARITY).
FT DOMAIN 517 527 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 161 164 POLY-LEU.
FT DOMAIN 1135 1141 POLY-ALA.
FT DOMAIN 1249 1252 POLY-SER.
FT CONFLICT 851 851 Q -> H (IN REF. 2).
FT CONFLICT 1180 1180 T -> S (IN REF. 2).
FT CONFLICT 1223 1223 S -> A (IN REF. 2).
SQ SEQUENCE 1343 AA; 150942 MW; 3FC2F286AC535F5 CRC64;

Query Match 42.5%; Score 45; DB 1; Length 1343;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 DSGSGKKITICE 13
Db 288 DNGSGGKESCE 299
RESULT 13
XEN3_XENLA STANDARD; PRT; 66 AA.
AC P38952;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE XENOXIN-3.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RN SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=93346392; PubMed=8393864;
RA Kolbe H.V.J., Huber A., Cordier P., Rasmussen U.B., Bouchon B.,
RA Jaquinod M., Vlasak R., Delot E.C., Kreil G.;
RT "Xenoxins, a family of peptides from dorsal gland secretion of
Xenopus laevis related to snake venom cytotoxins and neurotoxins.";
RL J. Biol. Chem. 268:16458-16464(1993).
CC -!- FUNCTION: LACKS ALPHA-NEUROTOXIC ACTIVITY, HAS APPARENTLY NO
ANTIBACTERIAL ACTIVITY, NOR ANTI-COAGULANT POTENCY.
CC -!- TISSUE SPECIFICITY: SKIN.
CC -!- SIMILARITY: TO SNAKE VENOM CYTOTOXINS AND NEUROTOXINS.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam: PF00087; toxin; 1.
KW Toxin.
FT DISULFID 3 24 BY SIMILARITY.
FT DISULFID 17 37 BY SIMILARITY.
FT DISULFID 43 58 BY SIMILARITY.
FT DISULFID 59 64 BY SIMILARITY.
FT UNSURE 23 35
FT UNSURE 42 47
SQ SEQUENCE 66 AA; 7259 MW; 36A2C26D981D88F5 CRC64;
Query Match 41.5%; Score 44; DB 1; Length 66;
Best Local Similarity 53.3%; Pred. No. 2.4;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 5 SNGKKITCECTKPD 19
Db 8 ANGKMTQCEAKEDT 22
RESULT 14
VEL_HPV70 STANDARD; PRT; 652 AA.
AC P50765;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE REPLICATION PROTEIN E1.
GN E1.
OS Human papillomavirus type 70.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=39457;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=96249586; PubMed=8815087;
RA Forslund O., Hansson B.G.;
RT "Human papillomavirus type 70 genome cloned from overlapping PCR
products: complete nucleotide sequence and genomic organization.";
RL J. Clin. Microbiol. 34:802-809(1996).

-1- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR INITIATION OF VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2 PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH CONTAINS BINDING SITES FOR BOTH PROTEINS.

-1- SUBCELLULAR LOCATION: NUCLEAR.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

EMBL: U21941; AAC54852.1; -

InterPro: IPR001177; Papillom_E1.

Pfam: PF00519; E1; 1.

Pfam: PF00524; E1_N; 1.

Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;

Nuclear protein.

NP_BIND 479 486 ATP (POTENTIAL).

SEQUENCE 652 AA; 73256 MW; 35B34BE9555739AB CRC64;

Query Match 41.5%; Score 44; DB 1; Length 652;
Best Local Similarity 47.4%; Pred. No. 21;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 1 EDSSNGKKTCTCKPDS 19

I: | | | | | | | | | |

Db 162 ENGGENGSIRECGSVDS 180

RESULT 15

FOLG_TMEVD

ID POLG_TMEVD STANDARD; PRT; 2301 AA.
AC P33899; Q88564; Q88565; Q88566; Q88567; Q88568; Q88569; Q88570;
AC Q88571; Q88572; Q88573; Q88574; Q89580;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEINS
DE P2A TO P2C, P3A; GENOME-LINKED PROTEIN VPG; PICORNAIN 3C
DE (EC 3.4.22.28) (PROTEASE 3C) (P3C); RNA-DIRECTED RNA POLYMERASE P3D
DE (EC 2.7.7.48)].
OS Theiler's murine encephalomyelitis virus (strain DA) (TMEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Cardiovirus.
OX NCBI_TaxID=12126;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88206072; PubMed=2834872;
RA Ohara Y., Stein S., Fu J., Stillman L., Klamon L., Roos R.P.;
RT "Molecular cloning and sequence determination of DA strain of
RT Theiler's murine encephalomyelitis viruses.";
RL Virology 164:245-253(1988).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE=92196057; PubMed=1549565;
RA Grant R.A., Filman D.J., Fujinami R.S., Icenogle J.P., Hogle J.M.;
RT "Three-dimensional structure of Theiler virus";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2061-2065(1992).
CC -1- FUNCTION: IT IS THOUGHT THAT THE P2C PROTEIN ATTACHES TO VESICULAR
CC MEMBRANES AND IS ASSOCIATED WITH VIRAL RNA SYNTHESIS.
CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS.
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: CLOSELY RELATED TO ENCEPHALOMYOCARDITIS VIRUS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

EMBL: M20301; AAA47928.1; -

PIR: A31228; GNNVTN.

PDB: 1TME; 31-JAN-94.

DR MEROPS; C03.009; -

DR MEROPS; U29.001; -

DR InterPro: IPR000605; RNA_helicase.

DR InterPro: IPR001205; RNA_pol_P3D.

DR InterPro: IPR001676; Rhv.

DR Pfam; PF00073; rhv; 3.

DR Pfam; PF00580; RNA_dep_RNA_pol; 1.

DR Pfam; PF00910; RNA_helicase; 1.

KW Polyprotein; Coat protein; Core protein; Transferase;

KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;

3D-structure.

FT PROPEP 1 76 LEADER PEPTIDE.

FT CHAIN 77 147 COAT PROTEIN VP4 (P1A).

FT CHAIN 148 414 COAT PROTEIN VP2 (P1B).

FT CHAIN 415 646 COAT PROTEIN VP3 (P1C).

FT CHAIN 647 920 COAT PROTEIN VP1 (P1D).

FT CHAIN 921 1062 CORE PROTEIN P2A.

FT CHAIN 1063 1189 CORE PROTEIN P2B.

FT CHAIN 1190 1515 CORE PROTEIN P2C.

FT CHAIN 1516 1603 CORE PROTEIN P3A.

FT CHAIN 1604 1623 GENOME-LINKED PROTEIN VPG (P3B).

FT CHAIN 1624 1840 PICORNAIN 3C.

FT CHAIN 1841 2301 RNA-DIRECTED RNA POLYMERASE P3D.

FT LIPID 77 77 MYRISTATE (BY SIMILARITY).

FT ACT_SITE 1791 1791 PROTEASE (POTENTIAL).

FT ACT_SITE 1809 1809 PROTEASE (POTENTIAL).

FT TURN 96 97

FT HELIX 103 106

FT STRAND 109 110

FT STRAND 162 166

FT TURN 167 168

FT STRAND 169 173

FT STRAND 179 181

FT HELIX 182 184

FT TURN 193 194

FT STRAND 200 201

FT HELIX 204 206

FT STRAND 210 217

FT TURN 219 220

FT TURN 223 224

FT STRAND 226 231

FT TURN 232 232

FT HELIX 233 235

FT HELIX 238 247

FT TURN 248 249

FT STRAND 250 262

FT TURN 267 268

FT STRAND 270 279

FT TURN 280 280

FT STRAND 286 287

FT TURN 289 291

FT STRAND 294 295

FT TURN 297 298

FT STRAND 305 305

FT STRAND 324 324

FT TURN 325 326

FT HELIX 330 335

FT STRAND 338 342

FT TURN 343 345

FT STRAND 348 353

FT STRAND 362 362

FT HELIX 364 366

FT STRAND 367
FT STRAND 370
FT TURN 381
FT STRAND 384
FT STRAND 385
FT STRAND 390
FT STRAND 406
FT TURN 423
FT TURN 430
FT STRAND 431
FT STRAND 438
FT STRAND 444
FT STRAND 445
FT TURN 449
FT STRAND 450
FT STRAND 454
FT STRAND 455
FT STRAND 457
FT STRAND 458
FT STRAND 463
FT STRAND 466
FT STRAND 470
FT STRAND 471
FT TURN 473
FT STRAND 474
FT STRAND 477
FT STRAND 478
FT STRAND 480
FT STRAND 484
FT STRAND 493
FT TURN 499
FT TURN 500
FT TURN 503
FT TURN 506
FT STRAND 513
FT TURN 514
FT TURN 515
FT STRAND 516
FT STRAND 520
FT STRAND 523
FT TURN 529
FT STRAND 533
FT STRAND 534
FT STRAND 544
FT STRAND 554
FT STRAND 557
FT TURN 558
FT STRAND 559
FT STRAND 561
FT STRAND 566
FT STRAND 572
FT STRAND 577
FT STRAND 586
FT STRAND 587
FT STRAND 601
FT STRAND 611
FT TURN 614
FT STRAND 615
FT STRAND 619
FT TURN 630
FT STRAND 631
FT STRAND 633
FT STRAND 637
FT STRAND 648
FT STRAND 651
FT STRAND 653
FT TURN 661
FT TURN 664
FT TURN 676
FT STRAND 677
FT STRAND 678
FT STRAND 683
FT STRAND 687
FT STRAND 688
FT STRAND 692
FT STRAND 698
FT STRAND 699
FT STRAND 706
FT STRAND 707
FT TURN 708
FT STRAND 709
FT STRAND 710
FT STRAND 711
FT STRAND 713
FT STRAND 715
FT STRAND 721
FT STRAND 722
FT TURN 727
FT TURN 728
FT TURN 732
FT STRAND 733
FT STRAND 737
FT STRAND 741
FT STRAND 744
FT STRAND 752
FT STRAND 755
FT STRAND 757
FT STRAND 758
FT STRAND 762
FT TURN 763
FT STRAND 766
FT STRAND 767
FT STRAND 786
FT STRAND 792
FT TURN 794
FT TURN 795
FT TURN 806
FT STRAND 807
FT STRAND 810
FT STRAND 812
FT STRAND 818
FT TURN 822
FT STRAND 824
FT STRAND 825
FT STRAND 828
FT STRAND 833
FT STRAND 842
FT STRAND 843
FT STRAND 847
FT STRAND 849
FT STRAND 853
FT TURN 854
FT TURN 859
FT STRAND 860
FT TURN 863

FT STRAND 869
FT STRAND 874
FT STRAND 878
SQ SEQUENCE 892
2301 AA; 256159 MM; OB6095DF153DBFDF CRC64;
Query Match 41.5%; Score 44; DB 1; Length 2301;
Best Local Similarity 70.0%; Pred. No. 67;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 6 NGKKITCCT 15
Db 1253 NGRKAYCCT 1262

Search completed: January 29, 2002, 11:13:43
Job time: 819 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:12:12 ; Search time 285.36 Seconds
(without alignments)
9.739 Million cell updates/sec

Title: US-09-763-397A-12

Perfect score: 106
Sequence: 1 EDGSGNGKKITCECTKPD5 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organalle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106	100.0	77	5 Q9BJV6	Q9bjv6 plasmodium
2	106	100.0	77	5 Q9BJV5	Q9bjv5 plasmodium
3	106	100.0	77	5 Q9BJV4	Q9bjv4 plasmodium
4	106	100.0	77	5 Q9BJV3	Q9bjv3 plasmodium
5	106	100.0	77	5 Q9BJV2	Q9bjv2 plasmodium
6	106	100.0	77	5 Q9BJV1	Q9bjv1 plasmodium
7	106	100.0	77	5 Q9BJV0	Q9bjv0 plasmodium
8	106	100.0	77	5 Q9BJU9	Q9bjv9 plasmodium
9	106	100.0	77	5 Q9BJU0	Q9bjv0 plasmodium
10	106	100.0	77	5 Q9BJT9	Q9bjt9 plasmodium
11	106	100.0	77	5 Q9BJT8	Q9bjt8 plasmodium
12	106	100.0	77	5 Q9BJT7	Q9bjt7 plasmodium
13	106	100.0	77	5 Q9BJT6	Q9bjt6 plasmodium
14	106	100.0	77	5 Q9BJT5	Q9bjt5 plasmodium
15	106	100.0	77	5 Q9BJT4	Q9bjt4 plasmodium
16	106	100.0	77	5 Q9BJT3	Q9bjt3 plasmodium
17	106	100.0	77	5 Q9BJT2	Q9bjt2 plasmodium
18	106	100.0	77	5 Q9BJT1	Q9bjt1 plasmodium
19	106	100.0	77	5 Q9BJT0	Q9bjt0 plasmodium

20	106	100.0	77	5 Q9BJS9	Q9bjs9 plasmodium
21	106	100.0	96	5 Q9TYE5	Q9tye5 plasmodium
22	106	100.0	116	5 Q9TYE3	Q9tye3 plasmodium
23	106	100.0	119	5 Q9TYE7	Q9tye7 plasmodium
24	106	100.0	121	5 Q9UAI6	Q9uai6 plasmodium
25	106	100.0	121	5 Q9TYE4	Q9tye4 plasmodium
26	106	100.0	219	5 Q9UB87	Q9ub87 plasmodium
27	106	100.0	372	5 Q43997	Q43997 plasmodium
28	106	100.0	372	5 Q25725	Q25725 plasmodium
29	106	100.0	372	5 Q25726	Q25726 plasmodium
30	106	100.0	373	5 Q25722	Q25722 plasmodium
31	106	100.0	373	5 Q25723	Q25723 plasmodium
32	106	100.0	373	5 Q25724	Q25724 plasmodium
33	106	100.0	376	5 Q9BMG8	Q9bmgs plasmodium
34	106	100.0	400	5 Q03999	Q03999 plasmodium
35	106	100.0	539	5 Q25966	Q25966 plasmodium
36	106	100.0	539	5 Q25971	Q25971 plasmodium
37	106	100.0	539	5 Q25972	Q25972 plasmodium
38	106	100.0	539	5 Q25973	Q25973 plasmodium
39	106	100.0	539	5 Q25976	Q25976 plasmodium
40	106	100.0	539	5 Q25981	Q25981 plasmodium
41	106	100.0	539	5 Q25984	Q25984 plasmodium
42	106	100.0	539	5 Q9TYG1	Q9tyg1 plasmodium
43	106	100.0	569	5 Q25967	Q25967 plasmodium
44	106	100.0	569	5 Q25969	Q25969 plasmodium
45	106	100.0	569	5 Q25974	Q25974 plasmodium

ALIGNMENTS

RESULT 1
Q9BJV6 PRELIMINARY; PRT; 77 AA.
AC Q9BJV6; 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
protein 1-19 and merozoite surface protein 2 in independent field
isolates of Plasmodium falciparum."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329507; AAK19326.1;
FT NON_TER 1 1
FT NON_TER 77 77
SQ SEQUENCE 77 AA; 8542 MW; 4EBC8D8B8AA169D4 CRC64;

Query Match 100.0%; Score 106; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 3e-10; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0;
QY 1 EDGSGNGKKITCECTKPD5 19
Db 54 EDGSGNGKKITCECTKPD5 72
|||||
RESULT 2
Q9BJV5 PRELIMINARY; PRT; 77 AA.
AC Q9BJV5; 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.

```
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329508; AAK19327.1; -
FT NON_TER 1
FT NON_TER 77
FT SEQUENCE 77 AA; 8542 MW; 4EBC8D8B8AA169D4 CRC64;
SQ

Query Match 100.0%; Score 106; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDGSGNGKKITCECTKPD 19
Db 54 EDGSGNGKKITCECTKPD 72

RESULT 3
Q9BJV4 PRELIMINARY; PRT; 77 AA.
AC Q9BJV4;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329509; AAK19328.1; -
FT NON_TER 1
FT NON_TER 77
FT SEQUENCE 77 AA; 8542 MW; 4EBC8D8B8AA169D4 CRC64;
SQ

Query Match 100.0%; Score 106; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDGSGNGKKITCECTKPD 19
Db 54 EDGSGNGKKITCECTKPD 72

RESULT 4
Q9BJV3 PRELIMINARY; PRT; 77 AA.
AC Q9BJV3;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
```

```
RT isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329510; AAK19329.1; -
FT NON_TER 1
FT NON_TER 77
FT SEQUENCE 77 AA; 8542 MW; 4EBC8D8B8AA169D4 CRC64;
SQ

Query Match 100.0%; Score 106; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDGSGNGKKITCECTKPD 19
Db 54 EDGSGNGKKITCECTKPD 72

RESULT 5
Q9BJV2 PRELIMINARY; PRT; 77 AA.
AC Q9BJV2;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329511; AAK19330.1; -
FT NON_TER 1
FT NON_TER 77
FT SEQUENCE 77 AA; 8542 MW; 4EBC8D8B8AA169D4 CRC64;
SQ

Query Match 100.0%; Score 106; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDGSGNGKKITCECTKPD 19
Db 54 EDGSGNGKKITCECTKPD 72

RESULT 6
Q9BJV1 PRELIMINARY; PRT; 77 AA.
AC Q9BJV1;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329512; AAK19331.1; -
FT NON_TER 1
FT NON_TER 77
FT SEQUENCE 77 AA; 8542 MW; 4EBC8D8B8AA169D4 CRC64;
SQ
```

Query Match 100.0%; Score 106; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPS 19
|||||
Db 54 EDGSGNGKKITCECTKPS 72

RESULT 7
Q9BJU0
ID Q9BJU0 PRELIMINARY; PRT; 77 AA.
AC Q9BJU0;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE MERZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329513; AAK19332.1; -
FT NON_TER 1
FT NON_TER 77
SQ SEQUENCE 77 AA; 8542 MW; 4EBC8D8B8AA169D4 CRC64;

Query Match 100.0%; Score 106; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPS 19
|||||
Db 54 EDGSGNGKKITCECTKPS 72

RESULT 8
Q9BJU9
ID Q9BJU9 PRELIMINARY; PRT; 77 AA.
AC Q9BJU9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE MERZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329514; AAK19333.1; -
FT NON_TER 1
FT NON_TER 77
SQ SEQUENCE 77 AA; 8542 MW; 4EBC8D8B8AA169D4 CRC64;

Query Match 100.0%; Score 106; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPS 19
|||||

Db 54 EDGSGNGKKITCECTKPS 72

RESULT 9
Q9BJU0
ID Q9BJU0 PRELIMINARY; PRT; 77 AA.
AC Q9BJU0;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE MERZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329523; AAK19342.1; -
FT NON_TER 1
FT NON_TER 77
SQ SEQUENCE 77 AA; 8575 MW; 4EBC923CEB03C90 CRC64;

Query Match 100.0%; Score 106; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPS 19
|||||
Db 54 EDGSGNGKKITCECTKPS 72

RESULT 10
Q9BJT9
ID Q9BJT9 PRELIMINARY; PRT; 77 AA.
AC Q9BJT9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE MERZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329524; AAK19343.1; -
FT NON_TER 1
FT NON_TER 77
SQ SEQUENCE 77 AA; 8575 MW; 4EBC923CEB03C90 CRC64;

Query Match 100.0%; Score 106; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPS 19
|||||
Db 54 EDGSGNGKKITCECTKPS 72

RESULT 11
Q9BJT8
ID Q9BJT8 PRELIMINARY; PRT; 77 AA.
AC Q9BJT8;

```
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN SEQUENCE FROM N.A.
RP Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
protein 1-19 and merozoite surface protein 2 in independent field
isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329525; AAK19344.1; -.
FT NON_TER 1
FT NON_TER 77
SQ SEQUENCE 77 AA; 8575 MW; 4EBC923ECEB03C90 CRC64;

Query Match 100.0%; Score 106; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPS 19
Db 54 EDGSGNGKKITCECTKPS 72

RESULT 12
Q9BJT7 PRELIMINARY; PRT; 77 AA.
AC Q9BJT7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN SEQUENCE FROM N.A.
RP Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
protein 1-19 and merozoite surface protein 2 in independent field
isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329526; AAK19345.1; -.
FT NON_TER 1
FT NON_TER 77
SQ SEQUENCE 77 AA; 8575 MW; 4EBC923ECEB03C90 CRC64;

Query Match 100.0%; Score 106; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPS 19
Db 54 EDGSGNGKKITCECTKPS 72

RESULT 13
Q9BJT6 PRELIMINARY; PRT; 77 AA.
AC Q9BJT6;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
```

```
RN SEQUENCE FROM N.A.
RP Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
protein 1-19 and merozoite surface protein 2 in independent field
isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329527; AAK19346.1; -.
FT NON_TER 1
FT NON_TER 77
SQ SEQUENCE 77 AA; 8575 MW; 4EBC923ECEB03C90 CRC64;

Query Match 100.0%; Score 106; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPS 19
Db 54 EDGSGNGKKITCECTKPS 72

RESULT 14
Q9BJT5 PRELIMINARY; PRT; 77 AA.
AC Q9BJT5;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN SEQUENCE FROM N.A.
RP Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
protein 1-19 and merozoite surface protein 2 in independent field
isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329528; AAK19347.1; -.
FT NON_TER 1
FT NON_TER 77
SQ SEQUENCE 77 AA; 8541 MW; 4EBC9CDECEB03C90 CRC64;

Query Match 100.0%; Score 106; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCECTKPS 19
Db 54 EDGSGNGKKITCECTKPS 72

RESULT 15
Q9BJT4 PRELIMINARY; PRT; 77 AA.
AC Q9BJT4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN SEQUENCE FROM N.A.
RP Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
protein 1-19 and merozoite surface protein 2 in independent field
isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
```


DR EMBL; AF329529; AAK19348.1; -.
FT NON_TER 1
FT NON_TER 77
SQ SEQUENCE 77 AA; 8541 MW; 4EBC9CDECEB03C90 CRC64;

Query Match 100.0%; Score 106; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDGSGNGKKITCECTKPDS 19
Db 54 EDGSGNGKKITCECTKPDS 72

Search completed: January 29, 2002, 11:12:12
Job time: 768 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:21:46 ; Search time 310.82 Seconds
(without alignments)
4.051 Million cell updates/sec

Title: US-09-763-397A-13

Perfect score: 88

Sequence: 1 GISYKEVLAKYKDDLE 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_1101:*

- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT:*
- 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	88	100.0	17	21	AA170289		Plasmodium falcipa
2	88	100.0	350	21	AA170278		Recombinant vaccin
3	88	100.0	375	22	AA183926		A major merozoite
4	45	51.1	12	12	AA14845		Protein associated
5	45	51.1	688	17	AA100087		RAS-related protei
6	45	51.1	688	20	AA149812		Human glioblastoma
7	45	51.1	688	21	AA120623		pML5 human gliobla
8	44	50.0	311	20	AA137149		Amino acid sequenc
9	44	50.0	410	20	AA171497		Zif protein with z
10	44	50.0	511	18	AA134560		Thermococcus chito
11	44	50.0	511	19	AA149864		Thermococcus GC74-

12	42	47.7	107	21	AAG22587	Zea mays protein f
13	42	47.7	172	16	AA172811	Gamma-IFN/glucagon
14	42	47.7	175	16	AA172812	Gamma-IFN/glucagon
15	42	47.7	280	21	AAG33132	Zea mays protein f
16	42	47.7	293	21	AAG33131	Zea mays protein f
17	42	47.7	296	22	AA125263	Human protein sequ
18	42	47.7	306	21	AAG33130	Zea mays protein f
19	42	47.7	1566	16	AA179643	Immunodominant ant
20	41.5	47.2	693	22	AA139578	Human polypeptide
21	41.5	47.2	726	22	AA11364	Human polypeptide
22	41	46.6	229	20	AA103681	Amino acid sequenc
23	41	46.6	262	22	AA137610	Merozoite surface
24	41	46.6	355	20	AA109372	Merozoite surface
25	41	46.6	355	20	AA105832	Merozoite surface
26	41	46.6	361	20	AA109373	Merozoite surface
27	41	46.6	361	20	AA105833	Merozoite surface
28	41	46.6	376	20	AA109374	Modified merozoite
29	41	46.6	376	20	AA105834	Modified merozoite
30	41	46.6	498	21	AA100038	LCMV coat glycopro
31	41	46.6	1639	19	AA154145	P. falciparum synt
32	41	46.6	1654	6	AA150777	Sequence of the p1
33	41	46.6	1921	21	AA154047	Arabidopsis thalia
34	41	46.6	1982	21	AA154046	Arabidopsis thalia
35	41	46.6	2003	21	AA154045	Arabidopsis thalia
36	40	45.5	164	21	AA14086	Arabidopsis thalia
37	40	45.5	164	21	AA160781	Arabidopsis thalia
38	40	45.5	171	21	AA149259	Arabidopsis thalia
39	40	45.5	201	21	AA14085	Arabidopsis thalia
40	40	45.5	206	21	AA149258	Arabidopsis thalia
41	40	45.5	239	22	AA100854	S.aureus conserved
42	40	45.5	264	20	AA132160	Soybean CDC-16 pro
43	40	45.5	298	21	AA14084	Arabidopsis thalia
44	40	45.5	303	21	AA149257	Arabidopsis thalia
45	40	45.5	481	18	AA134554	MILTL glycosidase

ALIGNMENTS

RESULT 1

AA170289
ID AA170289 standard; peptide: 17 AA.

AC AA170289;

DT 06-JUN-2000 (first entry)

XX Plasmodium falciparum MSP-1 antigenic epitope, p599.

DE Recombinant protein; CDC/NIAIDVAC-1; multivalent; malaria; vaccine;
KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
KW Circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
KW Liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
KW Apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
KW Pf927; antiparasitic; prevention; anti-CDC/NIAIDVAC-1 antibody.

OS Plasmodium falciparum.

PN WO200011179-A1.

XX 02-MAR-2000.

PF 19-AUG-1999; 99WO-US18869.

XX 21-AUG-1998; 98US-0097703.

XX (NIAID-) NAT INST IMMUNOLOGY.

PA (USHS) US DEPT HEALTH & HUMAN SERVICES.

PI Lal AA, Shi YP, Hasnain SE;

XX WPI; 2000-237654/20.

```

XX Novel recombinant protein as vaccine for treating malarial infection
PT comprises antigenic peptides obtained from different stages of
PT plasmodium falciparum life cycle
XX
XX Claim 2; Page 17; 52pp; English.
XX
XX The present sequence is the antigenic epitope p599, derived from
CC merozoite surface protein-1 (MSP-1) of the asexual blood stage of
CC Plasmodium falciparum. It is used in the construction of recombinant
CC protein CDC/NIIMALVAC-1, which is a multivalent, multistage malarial
CC vaccine. The recombinant protein comprises, melittin signal peptide,
CC (His)6 tag, T-cell epitope from tetanus toxoid and 21 antigenic epitopes
CC from circumsporozoite protein (CSP), sporozoite surface protein-2
CC (SSP-2), liver stage antigen-1 (LSA-1), merozoite surface protein-1
CC (MSP-1), MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding
CC antigen-175 (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete
CC specific antigen, Pf927. These epitopes were obtained at different stages
CC of the life cycle of P. falciparum. CDC/NIIMALVAC-1 vaccine has
CC antiparasitic activity and can be used for treatment and prevention of
CC malarial infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for
CC detecting P. falciparum in biological samples.
XX
XX Sequence 17 AA;
SQ
Query Match 100.0%; Score 88; DB 21; Length 17;
Best Local Similarity 100.0%; Pred. NO. 4e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GISYVEKVLAKYKDDLE 17
DB 1 gisyveklakykddle 17
RESULT 2
AAY70278
ID AAY70278 standard; Protein; 350 AA.
XX
XX AAY70278;
XX
XX 06-JUN-2000 (first entry)
XX
XX Recombinant vaccine CDC/NIIMALVAC-1.
XX
XX Recombinant protein: CDC/NIIMALVAC-1; multivalent; malaria; vaccine;
XX T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
XX circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
XX liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
XX apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
XX EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
XX Pf927; antiparasitic; prevention; anti-CDC/NIIMALVAC-1 antibody;
XX honey bee.
XX
XX Chimeric - Apis sp.
XX Chimeric - Clostridium tetani.
XX Chimeric - Plasmodium falciparum.
XX
XX Key Location/Qualifiers
XX Peptide 1..22
XX /label= Melittin_signal_peptide
XX /note= "Derived from Honey bee."
XX Protein 23..350
XX /label= Mature_CDC/NIIMALVAC-1
XX /note= "Recombinant multivalent malarial vaccine"
XX
XX WO200011179-A1.
XX
XX 02-MAR-2000.
XX
XX 19-AUG-1999; 99WO-US18869.
XX
XX 21-AUG-1998; 98US-0097703.

```

```

XX (NAlM-) NAT INST IMMUNOLOGY.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Lal AA, Shi YP, Hasnain SE;
XX
XX WPI; 2000-237654/20.
DR N-PSDB; AAZ51336.
XX
XX Novel recombinant protein as vaccine for treating malarial infection
PT comprises antigenic peptides obtained from different stages of
PT plasmodium falciparum life cycle
XX
XX Claim 3; Page 43-44; 52pp; English.
XX
XX The present sequence is that of recombinant protein CDC/NIIMALVAC-1,
CC which is a multivalent, multistage malarial vaccine. The recombinant
CC protein comprises, melittin signal peptide, (His)6 tag, T-cell epitope
CC from tetanus toxoid and 21 antigenic epitopes from circumsporozoite
CC protein (CSP), sporozoite surface protein-2 (SSP-2), liver stage
CC antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1), MSP-2, apical
CC membrane antigen-1 (AMA-1), erythrocyte binding antigen-175 (EBA-175),
CC rhoptry associated protein-1 (RAP-1) and gamete specific antigen, Pf927.
CC These epitopes were obtained at different stages of the life cycle of
CC Plasmodium falciparum. CDC/NIIMALVAC-1 vaccine has antiparasitic
CC activity and can be used for treatment and prevention of malarial
CC infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for detecting
CC P. falciparum in biological samples.
XX
XX Sequence 350 AA;
SQ
Query Match 100.0%; Score 88; DB 21; Length 350;
Best Local Similarity 100.0%; Pred. NO. 9.8e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GISYVEKVLAKYKDDLE 17
DB 334 gisyveklakykddle 350
RESULT 3
AAB83926
ID AAB83926 standard; Protein; 375 AA.
XX
XX AAB83926;
XX
XX 23-JUL-2001 (first entry)
XX
XX A major merozoite surface protein-1 fragment of 42kDa.
XX
XX Major merozoite surface protein-1; MSP1-42; melittin signal peptide;
XX malaria vaccine.
XX
XX Plasmodium falciparum.
XX
XX WO200134188-A1.
XX
XX 17-MAY-2001.
XX
XX 09-NOV-2000; 2000WO-US31064.
XX
XX 12-NOV-1999; 99US-0165178.
XX
XX 01-DEC-1999; 99US-0168327.
XX
XX 22-AUG-2000; 2000US-0226861.
XX
XX (UYHA-) UNIV HAWAII.
XX (UYCH-) UNIV CHINESE HONG KONG.
XX (QUEE-) QUEEN EMMA FOUND.
XX
XX Hul GSN, Lap-Yin P, Ho WKK;
XX
XX WPI; 2001-335879/35.

```

DR N-PSDB; AAF89840.
XX Producing malaria vaccine, useful for treatment or prevention of all
PT forms of malaria in humans, by expressing immunogenic merozoite protein
PT fragment in a baculovirus system -
XX
PS Example 3: Page 87-88; 95pp; English.
XX The present sequence represents a major merozoite surface protein-1
CC C-terminal fragment of 42kDa (MSPI-42). This fragment is linked to a
CC melittin signal peptide, and then expressed in a in a
CC silkworm/baculovirus system. The protein is used to prepare a
CC malaria vaccine, which is used to treat or prevent malaria, caused by
CC any of the four species of Plasmodium that infect humans.
XX
SQ Sequence 375 AA;

Query Match 100.0%; Score 88; DB 22; Length 375;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYVEKVLAKYKDDLE 17
Db 119 GISYVEKVLAKYKDDLE 135

RESULT 4
AAR1845
ID AAR1845 standard; Protein; 513 AA.
XX
AC AAR1845;
XX
XX 30-JAN-1992 (first entry)
XX
DE Protein associated with biochemical pathway involving CAMP.
XX
KW RAS; oncogene; cancer; cyclic nucleotide phosphodiesterase;
KW protein kinase.
XX
OS Homo sapiens.
XX
PN WO9116457-A.
XX
PD 31-OCT-1991.
XX
PF 19-APR-1991; 91WO-US02714.
XX
PR 20-APR-1990; 90US-0511715.
XX
PA (COLD-) COLD SPRING HARBOR.
XX
PI Wigler MH, Colicelli JJ;
XX
XX WPI; 1991-339841/46.
DR N-PSDB; AAQ14633.
XX
XX Complementary screening for genes and prods. - e.g. RAS protein
PT and CAMP, that modify, complement or suppress genetic defect and
PT correct associated phenotypic alteration
XX
PS Disclosure; Page 104; 169pp; English.
XX
XX In the specification this sequence is given the SEQ ID NO. 26 and
CC is encoded by the nucleotide sequence having SEQ ID NO.25.
CC References to SEQ ID NO. 25 in the text give the clone as S46 which
CC was selected by complementation in S.cerevisiae strain RS60.15B.
CC The yeast strain carries a mutant RAS2 allele such that cells are
CC unable to grow at 36 deg.C. There are two open reading frames and
CC the two corresponding amino acid sequences are printed as SEQ ID
CC NO.s 26 and 27, although only NO.26 is actually mentioned in the
CC text. It is described as being homologous to the S6 protein kinase
CC of Xenopus laevis. N.B. many of the references in the text of

CC the specification to particular sequences are incorrect, therefore
CC caution is advised in accepting that this sequence is indeed
CC an S6-like protein kinase !
XX
SQ Sequence 513 AA;

Query Match 51.1%; Score 45; DB 12; Length 513;
Best Local Similarity 41.2%; Pred. No. 58;
Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 GISYVEKVLAKYKDDLE 17
Db 80 gidyydrnlalfeeld 96

RESULT 5
AAW00087
ID AAW00087 standard; Protein; 688 AA.
XX
AC AAW00087;
XX
DT 09-OCT-1996 (first entry)
XX
DE RAS-related protein encoded by plasmid pML5 (ATCC 68593).
XX
KW Human; glioblastoma; complementation; S. cerevisiae; S. pombe;
KW Clone S46; strain RS60.15B; RAS2; RAS2(vall19ala15); Xenopus laevis;
KW S6 protein kinase; Plasmid pML5; strain SKN37; CAP; adenylyl cyclase;
KW diploid; strain SP565; ras1; ras1::LEU2; mutation.
XX
OS Homo sapiens.
XX
PN US5527896-A.
XX
PD 18-JUN-1996.
XX
PF 20-APR-1990; 90US-0511715.
XX
PR 19-APR-1991; 91US-0688352.
PR 20-APR-1990; 90US-0511715.
XX
XX (COLD-) COLD SPRING HARBOR LAB.
XX
XX Colicelli JJ, Wigler MH;
XX
DR WPI; 1996-299902/30.
DR N-PSDB; AAT34369.
XX
XX DNA mols. isolated from human glioblastoma cells - encode
PT RAS-related or cyclic nucleotide phosphodiesterase proteins
XX
PS Claim 1; Column 95-100; 101pp; English.
XX
XX The sequences given in AAW00087-89 and AAW00101 are encoded by plasmid
CC fragments which contain human glioblastoma cell cDNA inserts which do
CC not encode phosphodiesterases. The cDNA's were obtained by
CC complementation of two genetically altered S. cerevisiae and S. pombe
CC strains. Clone S46 was selected by complementation in S. cerevisiae
CC strain RS60.15B. This strain contains a mutant allele of RAS2,
CC RAS2(vall19ala15), which renders cells unable to grow at 36 deg. C,
CC because such cells are defective in RAS function at elevated
CC temperatures. Human cDNA's from a human glioblastoma cell library were
CC selected that could complement this defect. The deduced sequence of S46
CC is homologous to a Xenopus laevis gene that encodes a known protein
CC kinase, the S6 protein kinase. Plasmid pML5 was selected by
CC complementation in S. cerevisiae strain SKN37. This strain contains a
CC disrupted allele of CAP, cap::HIS3. CAP encodes an adenylyl cyclase
CC associated protein of undetermined function. As a consequence of this
CC gene disruption, SKN37 fails to grow in medium rich in amino acids.
CC Human cDNA's were chosen which could complement this defect. Plasmids
CC pATG16 and pATG29 were selected by complementation in the S. pombe
CC diploid strain SP565. This strain is homozygous for disruptions of ras1

CC (rasl::LEU2). As a result of this mutation, this strain fails to
CC sporulate, and human cDNA's were selected which could complement this
CC defect. These genes have unknown function.
XX
SQ Sequence 688 AA;

Query Match 51.1%; Score 45; DB 17; Length 688;
Best Local Similarity 41.2%; Pred. No. 79;
Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17
|||::||:::|:
Db 80 gidyydrnlalfceeld 96

RESULT 6
AAY49812
ID AAY49812 standard; Protein; 688 AA.
XX AC AAY49812;
XX
DT 19-JAN-2000 (first entry)
XX
DE Human glioblastoma cell RAS-related pML5 protein.
XX
KW phosphodiesterase; dunce-like phosphodiesterase; PDE; DPD; CAMP;
KW RAS-related protein; immunoreactive; detection; genetic defect;
KW bronchodilation; increased myocardial contractility;
KW anti-inflammation.
XX
OS Homo sapiens.
XX
FN US5977305-A.
XX
PD 02-NOV-1999.
XX
PF 07-JUN-1995; 95US-0474379.
XX
PR 01-MAR-1994; 94US-0206188.
PR 20-APR-1990; 90US-0511715.
PR 19-APR-1991; 91US-0688352.
XX
PA (COLD-) COLD SPRING HARBOR LAB.
XX
PI Colicelli JJ, Wigler MH;
XX
DR WPI; 1999-619709/53.
DR N-PSDB; AAZ32244.
XX
PT New isolated RAS-related polypeptides and mammalian cyclic nucleotide
PT phosphodiesterases, used for screening for agents which can modify
PT complement or suppress genetic defects
XX
PS Claim 1; Column 107-112; 145pp; English.
XX
CC The present invention describes new isolated RAS-related polypeptides
CC and mammalian cyclic nucleotide phosphodiesterases (PDEs). RAS-related
CC polypeptides are capable of complementing a defective RAS function in
CC yeast. The products can be used for screening for agents which can
CC modify, complement or suppress a genetic defect in a biochemical
CC pathway in which CAMP participates, or in a biochemical pathway which
CC is controlled, directly or indirectly, by a RAS protein and other
CC proteins affecting cell growth and maintenance. Developing agents that
CC will selectively act upon PDEs is directed toward reproducing the
CC desirable effects of cyclic nucleotides, e.g. bronchodilation,
CC increased myocardial contractility, anti-inflammation, yet without
CC causing the undesirable effects, e.g. increased heart rate or enhanced
CC lipolysis. The products can also be used for therapeutic, diagnostic
CC and prognostic uses. AAZ32229 to AAZ32285, and AAY49803 to AAY49830,
CC represent sequences used in the exemplification of the present
CC invention.
XX

QY 1 GISYKVLAKYKDDLE 17
|||::||:::|:
Db 80 gidyydrnlalfceeld 96

RESULT 7
AAB20623
ID AAB20623 standard; Protein; 688 AA.
XX AC AAB20623;
XX
DT 14-DEC-2000 (first entry)
XX
DE pML5 human glioblastoma cell insert protein sequence SEQ ID NO:28.
XX
KW Detection; mammalian gene; yeast; microorganism; identification;
KW phenotype; characteristic; dunce-like phosphodiesterase; PDE; RAS;
KW RAS-related protein; genetic defect; hybridisation; probe.
XX
OS Homo sapiens.
OS Saccharomyces cerevisiae.
XX
PN US6100025-A.
XX
PD 08-AUG-2000.
XX
PF 01-MAR-1994; 94US-0206188.
XX
PR 20-APR-1990; 90US-0511715.
PR 19-APR-1991; 91US-0688352.
XX
PA (COLD-) COLD SPRING HARBOR LAB.
XX
PI Colicelli JJ, Wigler MH;
XX
DR WPI; 2000-531664/48.
DR N-PSDB; AAB88179.
XX
PT Novel isolated DNA encoding a mammalian cyclic nucleotide
PT phosphodiesterase is present in plasmids pPDE46, pPDE43 or pPDE339 and
PT is used to modify a genetic defect in a biochemical pathway in which
PT CAMP participates
XX
PS Example 1; Column 113-118; 145pp; English.
XX
CC The present invention describes a purified and isolated DNA (I) which
CC encodes a mammalian cyclic nucleotide phosphodiesterase and is an insert
CC present in the plasmids pPDE46 (ATCC 69552), pPDE43 (ATCC 69551) or
CC pPDE339 (ATCC 69550). The DNA molecules are used to modify, complement
CC or suppress a genetic defect in a biochemical pathway in which CAMP
CC participates and are also used as hybridisation probes. The present
CC invention also describes methods for detecting mammalian genes encoding
CC proteins which can function in microorganisms, particularly yeast, to
CC modify, complement, or suppress a genetic defect associated with an
CC identifiable phenotypic alteration or characteristic in the
CC microorganism. AAB88162 to AAB88218 and AAB29614 to AAB20640 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 688 AA;

Query Match 51.1%; Score 45; DB 21; Length 688;
Best Local Similarity 41.2%; Pred. No. 79;
Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17
|||::||:::|:
Db 80 gidyydrnlalfceeld 96

Db 80 gidydrnlalfeceld 96
|||:|:|:|:|:

RESULT 8
AAV37149
ID AAV37149 standard; Protein; 311 AA.

XX AC AAV37149;

XX DT 07-OCT-1999 (first entry)

XX DE Amino acid sequence of a Chlamydia trachomatis protein.

XX KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; perihhepatitis;
KW nongonococcal urethritis; epidymitis; cervicitis; salpingitis;
KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.

XX OS Chlamydia trachomatis.

XX PN WO9928475-A2.

XX PD 10-JUN-1999.

XX PF 27-NOV-1998; 98WO-IB01939.

XX PR 04-NOV-1998; 98US-0107077.

XX PR 28-NOV-1997; 97FR-0015041.

XX PR 17-DEC-1997; 97FR-0016034.

XX PA (GEST) GENSET.

XX PI Griffais R;

XX DR WPI; 1999-371125/31.

XX PT Genome sequence of Chlamydia trachomatis

XX PS Disclosure; Page 933; 1755pp; English.

XX CC AAV36754-137949 are encoded by open reading frames (ORFs) of the genome
XX of Chlamydia trachomatis (see AAZ01425). The polypeptides can be used as
XX vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
XX can also be used to control growth of the microorganism. Chlamydia
XX trachomatis is responsible for a large number of diseases, e.g. eye
XX diseases such as conventional trachoma, nonendemic trachoma,
XX paratrachoma, and inclusion conjunctivitis; genital diseases such as
XX nongonococcal urethritis, epidymitis, cervicitis, salpingitis,
XX perihhepatitis, bartholinitis; pneumopathy in breast feeding infants;
XX CC and venereal lymphogranulomatosis. The polypeptides of the invention
XX may be of use in treating these diseases.

XX SQ Sequence 311 AA;

Query Match 50.0%; Score 44; DB 20; Length 311;
Best Local Similarity 64.3%; Pred. No. 48;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GISYKEVLAAYKD 14
|||:|:|:|:|:
Db 86 gisflekklambkd 99

RESULT 9

AAV17497
ID AAV17497 standard; Protein; 410 AA.

XX AC AAV17497;

XX DT 04-AUG-1999 (first entry)

XX

zif protein with zoocin A activity.

zif: zooA; zoocin A activity; immunity factor; protection;
bacteriolytic enzyme; antibacterial; Streptococcal; sore throat;
dental caries; food; nutraceutical; dairy product; confectionery;
chewing gum.

Streptococcus equi.

WO9926969-A1.

03-JUN-1999.

23-NOV-1998; 98WO-NZ00171.

21-NOV-1997; 97NZ-0329227.

(NZPA-) NEW ZEALAND PASTORAL AGRIC RES INST LTD.
(UYOT-) UNIV OTAGO.

Beatson SA, Simmonds RS;

WPI; 1999-347690/29.

N-PSDB; AAX76141.

zoocin A immunity factor useful in the treatment or prevention of

Streptococcal sore throat or dental caries

Claim 1: Page 34-35; 52pp; English.

The present sequence is a zif (zoocin A immunity factor) protein.
The DNA molecule or vector for expressing zif is used to protect
organisms susceptible to the bacteriolytic activity of zoocin A.
Non-pathogenic organisms expressing zif can be genetically modified to
express zoocin A activity without being at risk from the bacteriolytic
activity of zoocin A. The organisms are therefore resistant to zoocin A.
The non-pathogenic organisms that are zoocin A resistant are especially
food grade bacteria, such as Streptococcus gordonii. The transformed
non-pathogenic organisms are useful as antibacterial compositions,
suitable for human ingestion as part of a foodstuff, nutraceutical, a
dairy product, confectionery or chewing gum. The antibacterial
compositions are useful for preventing or inhibiting the growth of
undesirable organisms susceptible to zoocin A. Additionally the
antibacterial composition can be used to treat or prevent Streptococcal
sore throat or dental caries. The antibacterials can also be used to
treat animals, particularly horses that are infected with group C
Streptococci. GRAS (generally recognized as safe) organisms transformed
with the zif encoding polynucleotides produce zoocin A without
vulnerability to the activity of the enzyme itself. The resulting
organisms then have a wide range of applicability in antibacterials,
particularly foodstuffs.

Sequence 410 AA;

Query Match 50.0%; Score 44; DB 20; Length 410;
Best Local Similarity 47.1%; Pred. No. 65;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 GISYKEVLAAYKDDLE 17
|:||||:|:
Db 215 gleyyekffdafkdkse 231

RESULT 10

AAW34560
ID AAW34560 standard; Protein; 511 AA.

XX AC AAW34560;

XX DT 12-MAR-1998 (first entry)

XX DE Thermococcus chitonophagus glycosidase.

XX	Glycosidase; thermostable; textile; food processing; pharmaceutical;
KW	detergent; baking; industry; Thermococcus; Staphylothermus;
KW	Pyrococcus; glucose; soluble oligosaccharide.
XX	
XX	Thermococcus chitonophagus.
OS	
XX	WO9725417-A1.
XX	
XX	17-JUL-1997.
PD	
XX	10-JAN-1997; 97WO-US000092.
XX	
XX	13-SEP-1996; 96US-0712612.
PR	
XX	11-JAN-1996; 96US-0583787.
XX	
XX	(RECO-) RECOMBINANT BIOCATALYSIS INC.
PA	
XX	Bylina EJ, Lam DE, Mathur EJ, Swanson RV;
PI	
XX	WPI: 1997-372858/34.
DR	N-PSDB; AAT93684.
XX	
XX	New thermostable glycosidase(s) - from Thermococcus, Staphylothermus
PT	and Pyrococcus, used in the textile, food processing,
PT	pharmaceutical, detergent and baking industries
XX	
XX	Claim 4; Fig 7; 82pp; English.
PS	
XX	
XX	The present sequence represents glycosidase isolated from Thermococcus
CC	chitonophagus. The enzyme or its encoding nucleic acid sequence is used
CC	for generating glucose from soluble oligosaccharides. The enzyme can be
CC	used in the food processing, pharmaceutical, textile, detergent and
CC	baking industries. The enzyme is also used to treat lactose intolerance,
CC	as a diagnostic reporter molecule, in corn wet milling or in the fruit
CC	juice industry. The enzymes can be used to hydrolyse guar gum to remove
CC	non-reducing terminal mannose residues. The nucleic acids encoding the
CC	enzyme may be used to generate probes to identify similar sequences.
XX	
XX	Sequence 511 AA;
XX	

```

Query Match          50.0%; Score 44; DB 18; Length 511;
Best Local Similarity 37.5%; Pred. No. 82;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY      1 GISYVEKVLAKYKDDL 16
      ||::||::||::||::
Db       338 gvnnytrevkyyqgpm 353

RESULT 11
AAW49864
ID      AAW49864 standard; Protein; 511 AA.
XX
XX      AAW49864;
AC
XX
XX      21-DEC-1998 (first entry)
XX
XX      Thermococcus GC74-22G glycosidase.
DE
XX
XX      Glycosidase: GC74-22G; thermostable enzyme; oligosaccharide;
KW      glucose; sugar; baking; textile; detergent; beta-galactosidase.
KW
XX
XX      Thermococcus chitonophagus strain GC74-22G.
OS
XX
XX      Key      Location/Qualifiers
FH      Misc-difference 1
FT      /note= "encoded by TTG"
FT
XX
XX      WO9824799-A1.
XX
XX      11-JUN-1998.
XX
XX

```

XX 08-DEC-1997; 97WO-US22623.
XX
XX 10-OCT-1997; 97US-0949026.
XX
XX 06-DEC-1996; 96US-0056916.
XX
XX (DIVE-) DIVERSA CORP.
XX
XX Bylina EJ, Lam DE, Mathur EJ, Swanson RV;
XX
XX WPI; 1998-362407/31.
XX
XX N-PSDB; AAV36913.
XX
XX Glycosidase enzymes from organisms of the genera Staphylothermus,
XX
XX Pyrococcus and Thermococcus - for deriving sugar from
XX
XX oligosaccharides, useful in the e.g. food processing, textile or
XX
XX baking industries
XX
XX Claim 1; Fig 7a-b; 92pp; English.
XX
XX This is the amino acid sequence of glycosidase GC74-22G deduced
XX
XX from a polynucleotide (see AAV36913) of Thermococcus chitonophagus
XX
XX GC74 clone (22G) that grows optimally at 85 degC. The sequence
XX
XX shows 46% amino acid identity to a beta-galactosidase of
XX
XX Sulfolobus sulfataricus ATCC 49535/MT4. The invention provides 18
XX
XX polynucleotides (see AAV36907-24) coding for thermostable glycosidases
XX
XX (see AAV49858-75) having glucosidase, alpha-galactosidase,
XX
XX beta-galactosidase, beta-mannosidase, beta-mannanase, endoglucanase
XX
XX or pullulanase activity. Vectors and host cells are also claimed.
XX
XX A method is provided for producing the enzymes by recombinant
XX
XX techniques. A claimed method for generating glucose from soluble
XX
XX cell oligosaccharides comprises contacting a sample (selected from
XX
XX dairy products, fruit juice, detergent, textile, guar gum, animal
XX
XX feed, plant biomass or waste product) containing oligosaccharides
XX
XX (selected from maltose, cellobiose, lactose, sucrose, raffinose,
XX
XX starchose, verbascose, cellulose, starch, amylose, glycogen,
XX
XX disaccharides, polysaccharides and pullulan) with one of the
XX
XX claimed glycosidases such that glucose is produced.
XX
XX Sequence 511 AA;

```

Query Match          50.0%; Score 44; DB 19; Length 511;
Best Local Similarity 37.5%;
Matches 6; Conservative 6; Mismatches 0; Gaps 0;

Qy      1 GISYYEKVLAKYKDDL 16
          I::II:::II:::
Db      338 gynytrevvkyqdp 353

RESULT 12
AAG22587
ID      AAG22587 standard; Protein; 107 AA.
XX
AC      AAG22587;
XX
DT      17-OCT-2000 (first entry)
XX
DE      Zea mays protein fragment SEQ ID NO: 25574.
XX
KW      protein identification; signal transduction pathway; metabolic pathway;
KW      hybridisation assay; genetic mapping; gene expression control; promoter;
KW      termination sequence; corn.
XX
OS      Zea mays subsp. mays.
XX
PN      EP1033405-A2.
XX
PD      06-SEP-2000.
XX
PF      25-FEB-2000; 2000EP-0301439.
XX

```


PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134236.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.

PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1998; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 47.7%; Score 42; DB 21; Length 107;
Best Local Similarity 58.3%; Pred. No. 32;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 6 EKVLAKYKDDLE 17
I:: I:: I:: I:: I::
Db 92 eqikakfkddld 103

RESULT 13
AAR72811
ID AAR72811 standard; protein; 172 AA.

XX AC AAR72811;
XX DT 19-DEC-1995 (first entry)
XX DE Gamma-IFN/glucagon fusion protein.
XX KW Gamma-IFN/glucagon fusion protein; peptide isolation method;
KW granule forming protein; protease cleavage; E. coli;
KW gamma-interferon; recombinant production.

OS Homo sapiens.
XX JN JP07090000-A.
XX PD 04-APR-1995.
XX PF 22-SEP-1993; 93JP-0236732.
XX PR 22-SEP-1993; 93JP-0236732.
XX PA (SHIO) SHIONOGI & CO LTD.
XX DR WPI; 1995-167255/22.

PT A new method for the isolation of peptide(s) - comprises fusing the

PT peptide to a granule forming protein and cleaving the peptide from
PT the resulting fusion protein using a protease
XX Example 1; Pages 11-12; 17pp; Japanese.
XX AAR72811 is a gamma-IFN/glucagon fusion protein, it was used to
CC demonstrate a new method for the isolation of peptides. The
CC fusion protein is recombinantly produced in transformed E. coli,
CC the glucagon is then cleaved from the granule forming protein
CC gamma-IFN, using a protease. The glucagon can now be isolated
CC from the cell culture.
XX Sequence 172 AA;

Query Match 47.7%; Score 42; DB 16; Length 172;
Best Local Similarity 43.8%; Pred. No. 52;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ISYYEKVLAKYKDDLE 17.
I:: I:: I:: I:: I::
Db 50 vsfykfkfnfkddqe 65

RESULT 14
AAR72812
ID AAR72812 standard; protein; 175 AA.

XX AC AAR72812;
XX DT 19-DEC-1995 (first entry)
XX DE Gamma-IFN/glucagon/BNP fusion protein.
XX KW Gamma-IFN/glucagon/BNP; fusion protein; peptide isolation method;
KW granule forming protein; protease cleavage; E. coli;
KW B-type natriuretic peptide; gamma-interferon;
KW recombinant production.

XX OS Homo sapiens.
XX JN JP07090000-A.
XX PD 04-APR-1995.
XX PF 22-SEP-1993; 93JP-0236732.
XX PR 22-SEP-1993; 93JP-0236732.
XX PA (SHIO) SHIONOGI & CO LTD.
XX DR WPI; 1995-167255/22.

XX A new method for the isolation of peptide(s) - comprises fusing the
PT peptide to a granule forming protein and cleaving the peptide from
PT the resulting fusion protein using a protease
XX Example 2; Page 12; 17pp; Japanese.

XX AAR72812 is a gamma-IFN/glucagon/BNP fusion protein, it was used
CC to demonstrate a new method for the isolation of peptides. The
CC fusion protein is recombinantly produced in transformed E. coli,
CC the glucagon/BNP is then cleaved from the granule forming protein
CC gamma-IFN, using a protease. The glucagon can now be isolated
CC from the cell culture.

XX Sequence 175 AA;

Query Match 47.7%; Score 42; DB 16; Length 175;
Best Local Similarity 43.8%; Pred. No. 53;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Oy 2 ISYEKVLAKYKDDLE 17
:|:| |: :||| |
Db 50 vsfykfktnfkddqe 65

RESULT 15
AAG33132
ID AAG33132 standard; Protein; 280 AA.
XX AC AAG33132;
XX DT 18-OCT-2000 (first entry)
XX DE Zea mays protein fragment SEQ ID NO: 40099.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.
OS Zea mays subsp. mays.
PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

us-09-763-397a-13.rag

Job time: 422 sec

PR	16-AUG-1999;	99US-0149368
PR	17-AUG-1999;	99US-0149175
PR	18-AUG-1999;	99US-0149426
PR	20-AUG-1999;	99US-0149722
PR	20-AUG-1999;	99US-0149723
PR	20-AUG-1999;	99US-0149929
PR	20-AUG-1999;	99US-0149902
PR	23-AUG-1999;	99US-0149930
PR	25-AUG-1999;	99US-0150366
PR	26-AUG-1999;	99US-0150884
PR	27-AUG-1999;	99US-0151065
PR	27-AUG-1999;	99US-0151066
PR	27-AUG-1999;	99US-0151080
PR	27-AUG-1999;	99US-0151303
PR	30-AUG-1999;	99US-0151438
PR	31-AUG-1999;	99US-0151930
PR	01-SEP-1999;	99US-0152363
PR	07-SEP-1999;	99US-0153758
PR	10-SEP-1999;	99US-0153070
PR	13-SEP-1999;	99US-0153758
PR	15-SEP-1999;	99US-0154018
PR	16-SEP-1999;	99US-0154039
PR	20-SEP-1999;	99US-0154779
PR	22-SEP-1999;	99US-0155339
PR	23-SEP-1999;	99US-0155486
PR	24-SEP-1999;	99US-0156458
PR	28-SEP-1999;	99US-0156559
PR	29-SEP-1999;	99US-0156596
PR	04-OCT-1999;	99US-0157117
PR	05-OCT-1999;	99US-0157753
PR	06-OCT-1999;	99US-0157865
PR	07-OCT-1999;	99US-0158029
PR	08-OCT-1999;	99US-0158232
PR	12-OCT-1999;	99US-0158369
PR	13-OCT-1999;	99US-0159293
PR	13-OCT-1999;	99US-0159294
PR	13-OCT-1999;	99US-0159295
PR	14-OCT-1999;	99US-0159329
PR	14-OCT-1999;	99US-0159330
PR	14-OCT-1999;	99US-0159331
PR	14-OCT-1999;	99US-0159637
PR	14-OCT-1999;	99US-0159638
PR	18-OCT-1999;	99US-0159584
PR	21-OCT-1999;	99US-0160741
PR	21-OCT-1999;	99US-0160767
PR	21-OCT-1999;	99US-0160768
PR	21-OCT-1999;	99US-0160770
PR	21-OCT-1999;	99US-0160814
PR	21-OCT-1999;	99US-0161406
PR	25-OCT-1999;	99US-0161359
PR	26-OCT-1999;	99US-0161360
PR	26-OCT-1999;	99US-0161361
PR	26-OCT-1999;	99US-0160981
PR	22-OCT-1999;	99US-0160989
PR	22-OCT-1999;	99US-0161404
PR	25-OCT-1999;	99US-0161405
PR	25-OCT-1999;	99US-0161406
PR	25-OCT-1999;	99US-0161599
PR	26-OCT-1999;	99US-0160810
PR	22-OCT-1999;	99US-0160980
PR	22-OCT-1999;	99US-0160981
PR	22-OCT-1999;	99US-0160989
PR	25-OCT-1999;	99US-0161920
PR	28-OCT-1999;	99US-0161992
PR	28-OCT-1999;	99US-0161993
PR	29-OCT-1999;	99US-0162142

```
Query Match          47.7%; Score 42; DB 21; Length 280;
Best Local Similarity 58.3%; Pred. NO. 88;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      6 EKVLAKYKDDLE 17
         |::|::|::|::|:
Db     265 eqikakfkddld 276
```

Search completed: January 29, 2002, 10:21:46

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:24:05 ; Search time 133.18 seconds
(without alignments)
2.872 Million cell updates/sec

Title: US-09-763-397A-13

Perfect score: 88

Sequence: 1 GISYKVKVLAKYKDDLE 17

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_AA.*
- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
 - 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
 - 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
 - 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
 - 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
 - 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	51.1	513	5	PCT-US91-02714-26
2	45	51.1	688	2	Sequence 26, Appl
3	45	51.1	688	1	Sequence 28, Appl
4	45	51.1	688	3	Sequence 28, Appl
5	45	51.1	688	3	Sequence 28, Appl
6	44	50.0	450	4	Sequence 28, Appl
7	42	47.7	1565	2	Sequence 13, Appl
8	40	45.5	105	6	Sequence 23, Appl
9	39	44.3	18	3	Patent No. 5182210
10	39	44.3	18	3	Sequence 238, App
11	39	44.3	18	3	Sequence 248, App
12	39	44.3	18	3	Sequence 248, App
13	39	44.3	18	3	Sequence 248, App
14	39	44.3	18	3	Sequence 248, App
15	39	44.3	18	4	Sequence 248, App
16	39	44.3	18	4	Sequence 248, App
17	39	44.3	18	6	Sequence 248, App
18	39	44.3	19	6	Patent No. 5182364
19	39	44.3	48	6	Patent No. 5168045
20	39	44.3	48	6	Patent No. 5168045
21	39	44.3	122	1	Patent No. 5182364
22	39	44.3	126	6	Sequence 9, Appl
23	39	44.3	127	6	Patent No. 5518899
24	39	44.3	128	6	Patent No. 5518899
25	39	44.3	129	6	Patent No. 5518899
26	39	44.3	138	6	Patent No. 5518899
27	39	44.3	143	1	Sequence 8, Appl

28	39	44.3	143	3	US-09-029-819-1	Sequence 1, Appl
29	39	44.3	143	3	US-09-029-819-3	Sequence 3, Appl
30	39	44.3	143	3	US-09-102-150-8	Sequence 8, Appl
31	39	44.3	144	3	US-09-029-819-2	Sequence 2, Appl
32	39	44.3	146	1	US-08-030-077-1	Sequence 1, Appl
33	39	44.3	146	6	5210029-4	Patent No. 5210029
34	39	44.3	146	6	5278286-3	Patent No. 5278286
35	39	44.3	147	6	5278286-2	Patent No. 5278286
36	39	44.3	166	3	US-08-765-381-8	Sequence 8, Appl
37	39	44.3	166	3	US-08-765-381-14	Sequence 14, Appl
38	39	44.3	166	4	US-09-206-935-24	Sequence 24, Appl
39	39	44.3	639	2	US-08-282-197C-56	Sequence 56, Appl
40	39	44.3	639	2	US-08-557-309B-37	Sequence 37, Appl
41	39	44.3	639	3	US-08-834-306-37	Sequence 37, Appl
42	39	44.3	639	4	US-08-993-674A-37	Sequence 37, Appl
43	38	43.2	270	2	US-08-436-748-3	Sequence 3, Appl
44	38	43.2	270	3	US-08-483-857-4	Sequence 4, Appl
45	38	43.2	289	4	US-08-961-083-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
PCT-US91-02714-26
; Sequence 26, Application PC/TUS9102714
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; NUMBER OF INVENTIONS: Processes
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02714
; FILING DATE: 19910419
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25447
; REFERENCE/DOCKET NUMBER: 27805/30197
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US91-02714-26

Query Match 51.1%; Score 45; DB 5; Length 513;
Best Local Similarity 41.2%; Pred. No. 19;
Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17
|| ||: || ||: ||:
Db 80 GIDYDRNLALFEELD 96

RESULT 2
US-07-688-352C-28
; Sequence 28, Application US/07688352C
; Patent No. 5527896
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/688,352C
; FILING DATE: 19910419
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/206,188
; FILING DATE: 01-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/688,352
; FILING DATE: 19-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 27866/32771
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 688 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-474-379C-28

Query Match 51.1%; Score 45; DB 2; Length 688;
Best Local Similarity 41.2%; Pred. No. 26;
Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17
|| ||: || ||: ||:
Db 80 GIDYDRNLALFEELD 96

RESULT 4
US-09-146-249A-28
; Sequence 28, Application US/09146249A
; Patent No. 6069240
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

QY 1 GISYKVLAKYKDDLE 17
|| ||: || ||: ||:
Db 80 GIDYDRNLALFEELD 96

RESULT 2
US-07-688-352C-28
; Sequence 28, Application US/07688352C
; Patent No. 5527896
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/688,352C
; FILING DATE: 19910419
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25447
; REFERENCE/DOCKET NUMBER: 27805/30197
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-5740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 688 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-688-352C-28

Query Match 51.1%; Score 45; DB 1; Length 688;
Best Local Similarity 41.2%; Pred. No. 26;
Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17
|| ||: || ||: ||:
Db 80 GIDYDRNLALFEELD 96

RESULT 3
US-08-474-379C-28
; Sequence 28, Application US/08474379C
; Patent No. 5977305
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED
; TITLE OF INVENTION: PROCESSES
; NUMBER OF SEQUENCES: 88

APPLICATION NUMBER: US/09/146,249A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: US 07/511,715
APPLICATION NUMBER: 20-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 688 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-146-249A-28

Query Match 51.1%; Score 45; DB 3; Length 688;
Best Local Similarity 41.2%; Pred. No. 26;
Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 GISYKVKVLAQKDDLE 17
|||::||:::|
DB 80 GIDYDRNLALFEELD 96

RESULT 5

US-08-206-188B-28
Sequence 28, Application US/08206188B
Patent No. 6100025

GENERAL INFORMATION:
APPLICANT: Wigler, Michael H.
APPLICANT: Colicelli, John J.
TITLE OF INVENTION: Cloning by Complementation and Related
NUMBER OF SEQUENCES: 84
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/206,188B
FILING DATE: 01-MAR-1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36107
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 688 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-206-188B-28

Query Match 51.1%; Score 45; DB 3; Length 688;
Best Local Similarity 41.2%; Pred. No. 26;
Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 GISYKVKVLAQKDDLE 17
|||::||:::|
DB 80 GIDYDRNLALFEELD 96

RESULT 6

US-09-306-593-13
Sequence 13, Application US/09306593
Patent No. 6184018

GENERAL INFORMATION:
APPLICANT: Li, Xin-Liang
APPLICANT: Ljungdahl, Lars G.
APPLICANT: Chen, Huizhong
APPLICANT: Ximenes, Eduardo A.
TITLE OF INVENTION: Beta-glucosidase Coding Sequences and Protein from
TITLE OF INVENTION: Orpinomyces PC-2
FILE REFERENCE: 31-98us
CURRENT APPLICATION NUMBER: US/09/306,593
EARLIER FILING DATE: 1999-05-06
EARLIER APPLICATION NUMBER: US 60/084,494
EARLIER FILING DATE: 1998-05-06
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 450
TYPE: PRT
ORGANISM: Thermoanaerobacter brockii

US-09-306-593-13

Query Match 50.0%; Score 44; DB 4; Length 450;
Best Local Similarity 40.0%; Pred. No. 24;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 GISYKVKVLAQKDD 15
|||::||:::|
DB 295 GVNYYTRSIKVD 309

RESULT 7

US-08-687-956A-23
Sequence 23, Application US/08687956A
Patent No. 5861157

GENERAL INFORMATION:
APPLICANT: BURNIE, JAMES P
APPLICANT: MATTHEWS, RUTH C
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
TITLE OF INVENTION: INFECTIONS DUE TO STREPTOCOCCI AND ENTEROCOCCI
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO, LLP
STREET: 1100 NEW YORK AVENUE, N.W., EAST TOWER, 9TH
STREET: FLOOR
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,956A
FILING DATE: 29-JUL-1996
CLASSIFICATION: 536

```

RESULT 10
US-08-940-095-248
; Sequence 248, Application US/089400095
; Patent No. 6004925
; GENERAL INFORMATION:
; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; APPLICANT: Dufourcq, Jean
; TITLE OF INVENTION: APOLIPOPROTEIN
; TITLE OF INVENTION: AND THEIR USE
; NUMBER OF SEQUENCES: 258
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

```

; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940,095
; FILING DATE: 29-SEP-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0004-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 248:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6004925e
; US-08-940-095-248

Query Match 44.3%; Score 39; DB 3; Length 18;
Best Local Similarity 58.3%; Pred. No. 5.5;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SYEKVLAKYKD 14
Db 5 AFYKVLKLEKE 16

RESULT 11
US-08-940-093-238
; Sequence 238, Application US/08940093
; Patent No. 6037323
; GENERAL INFORMATION:
; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
; NUMBER OF SEQUENCES: 258
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940,093
; FILING DATE: 29-SEP-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0006-999

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 238:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6037323e
; FEATURE:
; NAME/KEY: Other
; LOCATION: 1...18
; OTHER INFORMATION: N-terminal acetylated and
; OTHER INFORMATION: C-terminal amidated
; US-08-940-093-238

Query Match 44.3%; Score 39; DB 3; Length 18;
Best Local Similarity 58.3%; Pred. No. 5.5;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SYEKVLAKYKD 14
Db 5 AFYKVLKLEKE 16

RESULT 12
US-08-940-093-248
; Sequence 248, Application US/08940093
; Patent No. 6037323
; GENERAL INFORMATION:
; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
; NUMBER OF SEQUENCES: 258
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940,093
; FILING DATE: 29-SEP-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0006-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 248:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: No. 6037323e
US-08-940-093-248

Query Match 44.3%; Score 39; DB 3; Length 18;
Best Local Similarity 58.3%; Pred. No. 5.5;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SYYEKVLAKYKD 14
:||||| | |
Db 5 AFYEKVLKLEKE 16

RESULT 13

US-08-940-096-238
; Sequence 238, Application US/08940096
; Patent No. 6046166
; GENERAL INFORMATION:

; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther

; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS

; TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS

; NUMBER OF SEQUENCES: 258

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds LLP

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10036-2811

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/940,096

; FILING DATE: 29-SEP-1997

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A

; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 009196-0005-999

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-493-4935

; TELEFAX: 650-493-5556

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 238:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 18 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: No. 6046166e

; FEATURE:

; NAME/KEY: Other

; LOCATION: 1...18

; OTHER INFORMATION: N-terminal acetylated and

; OTHER INFORMATION: C-terminal amidated

US-08-940-096-238

Query Match 44.3%; Score 39; DB 3; Length 18;
Best Local Similarity 58.3%; Pred. No. 5.5;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SYYEKVLAKYKD 14
:||||| | |
Db 5 AFYEKVLKLEKE 16

RESULT 14

US-08-940-096-248
; Sequence 248, Application US/08940096
; Patent No. 6046166
; GENERAL INFORMATION:

; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther

; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS

; TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS

; NUMBER OF SEQUENCES: 258

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds LLP

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10036-2811

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/940,096

; FILING DATE: 29-SEP-1997

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A

; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 009196-0005-999

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-493-4935

; TELEFAX: 650-493-5556

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 248:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 18 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: No. 6046166e

US-08-940-096-248

Query Match 44.3%; Score 39; DB 3; Length 18;
Best Local Similarity 58.3%; Pred. No. 5.5;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SYYEKVLAKYKD 14
:||||| | |
Db 5 AFYEKVLKLEKE 16

RESULT 15

US-09-465-719-238
; Sequence 238, Application US/09465719
; Patent No. 6265377
; GENERAL INFORMATION:

; APPLICANT: Dasseux, Jean-Louis

; APPLICANT: Sekul, Renate

; APPLICANT: Buttner, Klaus

; APPLICANT: Cornut, Isabelle

```

; APPLICANT: Metz, Gunther
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
; TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
; NUMBER OF SEQUENCES: 258
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/465,719
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/940,093
; FILING DATE: 29-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0006-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 238:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6265377e
; FEATURE:
; NAME/KEY: Other
; LOCATION: 1...18
; OTHER INFORMATION: N-terminal acetylated and
; OTHER INFORMATION: C-terminal amidated
US-09-465-719-238

```

```

Query Match      44.3%; Score 39; DB 4; Length 18;
Best Local Similarity 58.3%; Pred. NO. 5.5;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 3 SYEKVLAKYKD 14
DB 5 AFYKVLKLE 16

```

```

Search completed: January 29, 2002, 10:24:06
Job time: 512 sec

```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:26:41 ; Search time 144.96 Seconds
(without alignments)
8.933 Million cell updates/sec

Title: US-09-763-397A-13
Perfect score: 88
Sequence: 1 GISYKVLAKYKDDLE 17
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	88	100.0	651	2	A54498	merozoite surface
2	88	100.0	1701	2	A54498	major merozoite su
3	88	100.0	1701	2	A26868	major merozoite su
4	88	100.0	1726	2	SAZQGM	major merozoite su
5	88	100.0	1726	2	A45948	major merozoite su
6	50	56.8	815	2	F86311	hypothetical prote
7	47	53.4	806	2	T25614	hypothetical prote
8	44.5	50.6	431	2	E45734	histidinol dehydro
9	44.5	50.6	431	2	A86776	histidinol dehydro
10	44	50.0	451	2	G71561	probable integral
11	44	50.0	540	2	A82083	conserved hypothet
12	44	50.0	587	2	A49942	hypothetical prote
13	44	50.0	1755	2	T51532	hypothetical prote
14	43.5	49.4	447	2	C83890	beta-glucosidase b
15	43	48.9	445	2	A71699	UDP-n-acetylmuramo
16	43	48.9	608	2	B82938	zinc metalloprotei
17	43	48.9	701	2	T07943	probable AMP-bindi
18	43	48.9	878	2	T43767	DNA topoisomerase
19	42	47.7	176	2	T17935	hypothetical prote
20	42	47.7	308	2	A83875	positive regulator
21	42	47.7	401	2	T25031	hypothetical prote
22	42	47.7	425	2	T19415	hypothetical prote
23	42	47.7	459	2	F83602	hypothetical prote
24	42	47.7	482	2	G64771	yaJK protein - Esc
25	42	47.7	482	2	T46944	thiamin biosynthes
26	42	47.7	530	2	T13090	probable minor cap
27	42	47.7	533	1	VHBPBL	minor capsid prote
28	42	47.7	533	2	B85690	hypothetical prote
29	42	47.7	636	2	C70031	hypothetical prote

30	42	47.7	911	2	T08105	nitrate reductase
31	42	47.7	928	2	C37271	A-alpha y 4 protei
32	42	47.7	1214	2	H75034	reverse gyrase (to
33	42	47.7	1528	2	A60338	surface antigen A
34	42	47.7	1566	2	A43607	cell surface anti
35	42	47.7	1624	2	C71129	probable reverse g
36	41.5	47.2	679	2	S38011	hypothetical prote
37	41	46.6	152	2	S76079	hypothetical prote
38	41	46.6	155	2	G82504	hypothetical prote
39	41	46.6	307	2	A75020	hypothetical prote
40	41	46.6	307	2	E71206	hypothetical prote
41	41	46.6	400	2	A45545	major merozoite su
42	41	46.6	401	2	T03118	hypothetical prote
43	41	46.6	498	1	VGXPLA	surface glycoprote
44	41	46.6	498	1	VGXPLA	surface glycoprote
45	41	46.6	559	2	A45620	cytoovillin homolog

ALIGNMENTS

RESULT 1

S47282
merozoite surface antigen 1 - malaria parasite (Plasmodium falciparum) (strain RO-71)
C:Species: Plasmodium falciparum
A:Variety: strain RO-71
C:Date: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 09-Jun-2000
C:Accession: S47282
R:Tolle, R.; Bujard, H.; Cooper, J.A.
submitted to the EMBL Data Library, July 1994
A:Description: Plasmodium falciparum: recombination within the C-terminal region of m
A:Reference number: S47282
A:Accession: S47282
A:Molecule type: DNA
A:Residues: 1-651 <TOL>
A:Cross-references: EMBL:Z35329; NID:g535257; PIDN:CAA84558.1; PID:g53525758
A:Experimental source: strain RO-71
C:Superfamily: major merozoite surface antigen
C:Keywords: glycoprotein; merozoite; surface antigen

Query Match 100.0%; Score 88; DB 2; Length 651;
Best Local Similarity 100.0%; Pred. No. 6.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GISYKVLAKYKDDLE 17
|||||
Db 376 GISYKVLAKYKDDLE 392

RESULT 2

A54498
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C:Accession: A54498
R:Peterson, M.G.; Coppel, R.L.; McIntyre, P.; Langford, C.J.; Woodrow, G.; Brown, G.V
Mol. Biochem. Parasitol. 27, 291-302, 1988
A:Title: Variation in the precursor to the major merozoite surface antigens of Plasmo
A:Reference number: A54498; MUID:88142999
A:Accession: A54498
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1701 <PPT>
A:Cross-references: GB:MJ9143; NID:g160412; PIDN:AAA29653.1; PID:g160413
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen

Query Match 100.0%; Score 88; DB 2; Length 1701;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17
|||||
Db 1425 GISYKVLAKYKDDLE 1441

RESULT 3
A26868
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (st
C:Species: Plasmodium falciparum
C:Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 09-Jun-2000
C:Accession: A26868
R:Tanabe, K.; Mackay, M.; Goman, M.; Scaife, J.G.
J. Mol. Biol. 195, 273-287, 1987
A:Title: Allelic dimorphism in a surface antigen gene of the malaria parasite Plasmodium
A:Reference number: A26868; MUID:88011243
A:Accession: A26868
A:Molecule type: DNA
A:Residues: 1-1701 <TAN>
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1701/Product: major merozoite surface antigen #status predicted <MAT>

Query Match 100.0%; Score 88; DB 2; Length 1701;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17
|||||
Db 1425 GISYKVLAKYKDDLE 1441

RESULT 4
SA206M
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (st
N:Alternate names: 195K glycoprotein
C:Species: Plasmodium falciparum
C:Date: 30-Sep-1987 #sequence_revision 31-Mar-1991 #text_change 09-Jun-2000
C:Accession: A23386; S06361
R:Weber, J.L.; Leininger, W.M.; Lyon, J.A.
Nucleic Acids Res. 14, 3311-3323, 1986
A:Title: Variation in the gene encoding a major merozoite surface antigen of the human m
A:Reference number: A23386; MUID:86205236
A:Accession: A23386
A:Molecule type: DNA
A:Residues: 1-1104 <WEB1>
A:Cross-references: EMBL:X03831
R:Weber, J.L.; Sim, B.K.L.; Lyon, J.A.; Wolff, R.
Nucleic Acids Res. 16, 1206, 1988
A:Title: Merozoite surface protein sequence from the Camp strain of the human malaria pa
A:Reference number: S06361; MUID:88143999
A:Accession: S06361
A:Molecule type: DNA
A:Residues: 1104-1726 <WEB2>
A:Cross-references: EMBL:X03831
C:Comment: The merozoite stages of different strains have strain-specific surface antigen
C:Comment: P. falciparum has three stages: sporozoite, merozoite, and gametocyte. The me
C:Superfamily: major merozoite surface antigen
C:Keywords: glycoprotein; merozoite; surface antigen; tandem repeat
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1726/Product: major merozoite surface antigen #status predicted <MAT>
F:67-87,91-96,100-105,109-120/Region: 3-residue repeats (S-G-T)
F:757-765/Region: 3-residue repeats (T-E-E)
F:133,272,501,567,638,827,839,924,944,990,1016,1114,1221,1613,1658/Binding site: carbohy

Query Match 100.0%; Score 88; DB 1; Length 1726;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17
|||||

Db 1450 GISYKVLAKYKDDLE 1466

RESULT 5
A45948
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jun-2000
C:Accession: A45948
R:Chang, S.P.; Kramer, K.J.; Yamaga, K.M.; Kato, A.; Case, S.E.; Siddiqui, W.A.
Exp. Parasitol. 67, 1-11, 1988
A:Title: Plasmodium falciparum: gene structure and hydropathy profile of the major me
A:Reference number: A45948; MUID:89005525
A:Accession: A45948
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1726 <CHA>
A:Cross-references: GB:M37213
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen

Query Match 100.0%; Score 88; DB 2; Length 1726;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17
|||||
Db 1450 GISYKVLAKYKDDLE 1466

RESULT 6
F86311
hypothetical protein AAF99810.1 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F86311
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maith, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: F86311
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-815 <STO>
A:Cross-references: GB:AE005172; NID:g9802741; PIDN:AAF99810.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 56.8%; Score 50; DB 2; Length 815;
Best Local Similarity 76.9%; Pred. No. 7.2;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 SYEKVLAKYKDD 15
|||||
Db 748 SYEKVLAKYKDD 760

RESULT 7
T25614
hypothetical protein C37H5.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T25614
R:Davidson, S.; Gillam, B.

submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans cosmid C37H5.
A:Reference number: Z20058

A:Accession: T25614

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-806 <DAV>

A:Cross-references: EMBL:U08315; PIDN: AAB42369.1; GSPDB:GN00023; CESP:C37H5.5

A:Experimental source: strain Bristol N2; clone C37H5

C:Genetics:

A:Gene: CESP:C37H5.5

A:Map position: 5

A:Introns: 95/3; 141/1; 254/3; 342/1; 365/3; 672/3; 750/2

Query Match 53.4%; Score 47; DB 2; Length 806;

Best Local Similarity 69.2%; Pred. No. 21;

Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 YEKVLAKYKDDLE 17

I:|:| |:| |:| |

Db 767 YDKVESKYLDLLE 779

RESULT 8

E45734

histidinol dehydrogenase (EC 1.1.1.23) HisD [validated] - Lactococcus lactis subsp. lact

C:Species: Lactococcus lactis subsp. lactis

C:Date: 24-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 18-Aug-2000

C:Accession: E45734; B36890

R:Delorme, C.; Godon, J.J.; Ehrlich, S.D.; Renault, P.

J. Bacteriol. 174, 6571-6579, 1992

A:Title: Histidine biosynthesis genes in Lactococcus lactis subsp. lactis.

A:Reference number: A45734; MUID:93015709

A:Accession: E45734

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-431

A:Experimental source: strain NCDO2118 (prototrophic for histidine)

A:Note: sequence extracted from NCBI backbone (NCBIN:115814, NCBI:P:115826)

R:Delorme, C.; Godon, J.J.; Ehrlich, S.D.; Renault, P.

J. Bacteriol. 175, 4391-4399, 1993

A:Title: Gene inactivation in Lactococcus lactis: histidine biosynthesis.

A:Reference number: A36890; MUID:93322317

A:Accession: B36890

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-37, 'E', 39-42, 'S', 44-68, 'P', 70-73, 'T', 75-163, 'T', 165-244, 'E', 246-247, 'N', 24

A:Experimental source: strain Il1403 (auxotrophic for histidine)

C:Superfamily: histidinol dehydrogenase; histidinol dehydrogenase homology

C:Keywords: histidine biosynthesis; oxidoreductase

F:35-429/Domain: histidinol dehydrogenase homology <HID>

Query Match 50.6%; Score 44.5; DB 2; Length 431;

Best Local Similarity 62.5%; Pred. No. 27;

Matches 10; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

Qy 3 SYYEK-VLAKYKDDLE 17

I:|:| |:| |:| |

Db 394 SYYPKEVLAEFKEDVE 409

RESULT 9

A86776

histidinol dehydrogenase (EC 1.1.1.23) [imported] - Lactococcus lactis subsp. lactis (st

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 31-Mar-2001

C:Accession: A86776

R:Boilotin, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrli

Genome Res. in press, 2001

A:Title: The complete genome sequence of the lactic acid bacterium.

A:Reference number: A86625

A:Accession: A86776

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-431 <STO>

A:Cross-references: GB:AE005176; NID:gl2724178; PIDN:AAK05307.1; GSPDB:GN00146

A:Experimental source: strain Il1403

C:Genetics:

A:Gene: hisD

C:Superfamily: histidinol dehydrogenase; histidinol dehydrogenase homology

C:Keywords: Oxidoreductase

Query Match 50.6%; Score 44.5; DB 2; Length 431;

Best Local Similarity 62.5%; Pred. No. 27;

Matches 10; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

Qy 3 SYYEK-VLAKYKDDLE 17

I:|:| |:| |:| |

Db 394 SYYPKEVLAEFKEDVE 409

RESULT 10

G71561

probable integral membrane protein - Chlamydia trachomatis (serotype D, strain UW3/Cx

C:Species: Chlamydia trachomatis

C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999

C:Accession: G71561

R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche

Science 282, 754-759, 1998

A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t

A:Reference number: A71570; MUID:99000809

A:Accession: G71561

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-451 <ARN>

A:Cross-references: GB:AE001281; GB:AE001273; NID:g3328454; PIDN:AAC67660.1; PID:g332

A:Experimental source: serotype D, strain UW-3/Cx

C:Genetics:

A:Gene: ytgC

Query Match 50.0%; Score 44; DB 2; Length 451;

Best Local Similarity 64.3%; Pred. No. 34;

Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GISYVEKVLAKYKD 14

I:|:| |:| |:| |

Db 86 GISFLEKKLAMHKD 99

RESULT 11

A82083

conserved hypothetical protein VC2386 [imported] - Vibrio cholerae (strain N16961 ser

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: A82083

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers

I, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: A82083

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-540 <HEI>

A:Cross-references: GB:AB004309; GB:AE003852; NID:g9656954; PIDN:AAF95529.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC2386

A:Map position: 1

```
Query Match          50.0%; Score 44; DB 2; Length 540;
Best Local Similarity 50.0%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 ISYYEKVLAKYKDDLE 17
   :|||: ||| |
Db 216 VKYERKAKYSDKSE 231

RESULT 12
S49942
hypothetical protein Y11036w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein Y19905.12
C:Species: Saccharomyces cerevisiae
C:Date: 13-Jan-1995 #sequence_revision 20-Feb-1995 #text_change 05-Nov-1999
C:Accession: S49942
R:Odell, C.; Bowman, S.
submitted to the EMBL Data Library, December 1994
A:Reference number: S49931
A:Accession: S49942
A:Molecule type: DNA
A:Residues: 1-587 <ODE>
A:Cross-references: GB:247047; EMBL:246861; NID:g603997; PID:g763310; GSPDB:GN00009; MIPS
C:Genetics:
A:Gene: MIPS:Y11036w
A:Map position: 9L
C:Superfamily: fos/jun DNA-binding domain homology
F:420-460/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match          50.0%; Score 44; DB 2; Length 587;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 2 ISYYEKVLAKYK 13
   :|||: ||| |
Db 471 LNYEKLISKFK 482

RESULT 13
T51532
hypothetical protein T20K14_150 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C:Accession: T51532
R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A:Reference number: Z25394
A:Accession: T51532
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1755 <SAT>
A:Cross-references: EMBL:AL391143
A:Experimental source: cultivar Columbia; BAC clone T20K14
C:Genetics:
A:Map position: 5
A:Introns: 90/2; 127/1; 209/3; 323/3; 359/3; 397/1; 430/2; 452/3; 574/3; 634/3; 7
3; 1637/3; 1679/3
A:Note: T20K14_150

Query Match          50.0%; Score 44; DB 2; Length 1755;
Best Local Similarity 57.1%; Pred. No. 1.4e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GISYVEKVLAKYKD 14
   ||| | ||| |
Db 850 GIKYFEKVAERIKD 863

RESULT 14
C83890
beta-glucosidase bgIA [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 08-Dec-2000
C:Accession: C83890
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A:Reference number: A83650; MUID:20263314
A:Accession: C83890
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-447 <STO>
A:Cross-references: GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BA05642.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: bgIA
C:Superfamily: Agrobacterium beta-glucosidase

Query Match          49.4%; Score 43.5; DB 2; Length 447;
Best Local Similarity 45.5%; Pred. No. 40;
Matches 10; Conservative 4; Mismatches 3; Indels 5; Gaps 1;

QY 1 GISYVEKVLAKYKD-----DLE 17
   ||| | :|||: |||
Db 292 GINYTGSVARYKENEGFLDLE 313

RESULT 15
A71699
UDP-n-acetylmuramoylalanine--d-glutamate ligase (murD) RP410 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C:Accession: A71699
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark
Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499
A:Accession: A71699
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-445 <AND>
A:Cross-references: GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAAL4867.1; PID:g386
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: murD; RP410
C:Superfamily: UDP-N-acetylmuramate--alanine ligase

Query Match          48.9%; Score 43; DB 2; Length 445;
Best Local Similarity 52.4%; Pred. No. 48;
Matches 11; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

QY 1 GISYVEKVLAKYK-----DDLE 17
   ||| | ||| |
Db 18 GISYVEELKNKYDLIVYDDLE 38

Search completed: January 29, 2002, 10:26:42
Job time: 653 sec
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:13:43 ; Search time 80.65 Seconds
(without alignments)
7.728 Million cell updates/sec

Title: US-09-763-397A-13

Perfect score: 88

Sequence: 1 GISYVEKVLAKYKDDLE 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	88	100.0	1682	1 MSP1_PLAF3	P19598 plasmodium
2	88	100.0	1701	1 MSP1_PLAFF	P13819 plasmodium
3	88	100.0	1701	1 MSP1_PLAFM	P08569 plasmodium
4	88	100.0	1726	1 MSP1_PLAFC	P04934 plasmodium
5	88	100.0	1726	1 MSP1_PLAFP	P50495 plasmodium
6	44.5	50.6	431	1 HSLX_LACLA	Q02136 lactococcus
7	44	50.0	451	1 Y069_CHLTR	O84072 chlamydia t
8	44	50.0	587	1 YID6_YEAST	P40535 saccharomyc
9	43	48.9	445	1 MURD_RICPR	Q92dc2 rickettsia
10	42	47.7	194	1 YMP9_CABEL	Q10952 caenorhabdi
11	42	47.7	425	1 CAB1_CABEL	Q93249 caenorhabdi
12	42	47.7	482	1 THIL_ECOLI	P77718 escherichia
13	42	47.7	482	1 THIL_SALTY	P55913 salmonella
14	42	47.7	533	1 VNCB_LAMBD	P03710 bacterioph
15	42	47.7	911	1 NIA1_BRANA	P39867 brassica na
16	42	47.7	928	1 MAY4_SCHCO	P37935 schizophyll
17	42	47.7	1528	1 SPAA_STRDO	P21979 streptococ
18	41.5	47.2	527	1 MET3_CANAL	Q97872 candida alb
19	41.5	47.2	679	1 YKR9_YEAST	P34237 saccharomyc
20	41	46.6	289	1 VINT_BPL2	P42540 bacterioph
21	41	46.6	498	1 VGLY_LYCVA	P09991 lymphocytic
22	41	46.6	992	1 EVC_HUMAN	P57679 homo sapien
23	41	46.6	1630	1 MSP1_PLAFK	P04932 plasmodium
24	41	46.6	1639	1 MSP1_PLAFW	P04933 plasmodium
25	40.5	46.0	964	1 IF3A_YEAST	P38249 saccharomyc
26	40	45.5	489	1 NIFD_ALCFA	Q44045 alcaligenes
27	40	45.5	1465	1 MYM2_HUMAN	P54296 homo sapien
28	39.5	44.9	102	1 RPOL_AERPE	Q9ye21 aeropyrium p
29	39	44.3	147	1 PFDA_SULSO	P58179 sulfolobus
30	39	44.3	165	1 ING_CERTO	P42162 cercocebus
31	39	44.3	165	1 ING_WACMO	P42163 macaca mula
32	39	44.3	166	1 ING_HUMAN	P01579 homo sapien
33	39	44.3	192	1 Y678_AQUAE	O66906 aquifex aeo

ALIGNMENTS

```

RESULT 1
MSPI_PLAF3
ID MSP1_PLAF3 STANDARD: PRT: 1682 AA.
AC P19598; Q25921;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
DE (PMMSA) (P190).
GN MSP-1.
OS Plasmodium falciparum (isolate ro-33 / Ghana).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID:5834;
RN [1]
RP SEQUENCE OF 1-1061 FROM N.A.
RX MEDLINE=88166657; PubMed=3327688;
RA Certa U., Rotmann D., Matile H., Reber-Liste R.;
RT "A naturally occurring gene encoding the major surface antigen
precursor p190 of Plasmodium falciparum lacks tripeptide repeats.";
RL EMBO J. 6:4137-4142(1987).
RN [2]
RP SEQUENCE OF 1032-1682 FROM N.A.
RX MEDLINE=95354793; PubMed=7628566;
RA Tolle R., Bujard H., Cooper J.A.;
RT "Plasmodium falciparum: variations within the C-terminal region of
merozoite surface antigen-1.";
RL Exp. Parasitol. 81:47-54(1995).
CC -! SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
(POTENTIAL).
CC -! PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42
KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)
-----
CC EMBL; M35727; AAA29715.1; -
CC EMBL; Y00087; CAA68280.1; -
CC EMBL; Z35326; CAA84555.1; -
CC PIR; S06286; S06286.
CC InterPro: IPR000561; EGF-like.
CC Pfam; PF00008; EGF; 1.
CC Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
Transmembrane; GPI-anchor.
CC SIGNAL 1 19
CC CHAIN 20 1682
CC TRANSMEM 1666 1682
CC CARBOHYD 233 233
CC CARBOHYD 462 462
CC CARBOHYD 528 528
CC CARBOHYD 599 599
CC POTENTIAL.
CC MEROZOITE SURFACE PROTEIN 1.
CC MEMBRANE ANCHOR (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

34 39 44.3 281 1 VNS1_INBGL
35 39 44.3 334 1 Y798_METUA
36 39 44.3 387 1 CEXY_CLOSR
37 39 44.3 405 1 Y872_METUA
38 39 44.3 413 1 FENR_SYNY3
39 39 44.3 426 1 ENO_HELPJ
40 39 44.3 426 1 ENO_HELPY
41 39 44.3 429 1 ENO_THEMA
42 39 44.3 498 1 VGLY_LYCVW
43 39 44.3 607 1 HEMA_CDVAA
44 39 44.3 607 1 HEMA_CDVAA6
45 39 44.3 615 1 YJ13_SCHPO

```

```

P12593 Influenza b
Q58208 methanococc
P40942 clostridium
Q58282 methanococc
Q55318 synechocyst
Q9ams6 helicobacte
P48285 helicobacte
P42848 thermotoga
P07399 lymphocytic
O66001 canine dist
O66000 canine dist
O13681 schizosacch

```

```
FT CARBOHYD 785 785 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 881 881 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 901 901 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 947 947 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1178 1178 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1569 1569 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1682 AA; 192462 MW; C82A1E159948CAD6 CRC64;

Query Match 100.0%; Score 88; DB 1; Length 1682;
Best Local Similarity 100.0%; Pred. No. 6.8e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYVEKVLAKYKDDLE 17
| | | | | | | | | | | | | | | | | |
DB 1407 GISYVEKVLAKYKDDLE 1423

RESULT 2
MSP1_PLAFF STANDARD; PRT; 1701 AA.
AC P13819;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
DE (PMMSA).
GN MSP-1.
OS Plasmodium falciparum (isolate FC27 / Papua New Guinea).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5837;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88142999; PubMed=2449612;
RA Peterson M.G., Coppel R.L., McIntyre P., Langford C.J., Woodrow G.,
RA Brown G.V., Anders R.F., Kemp D.J.;
RT "Variation in the precursor to the major merozoite surface antigens
RT of Plasmodium falciparum."
RL Mol. Biochem. Parasitol. 27:291-302(1988).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
(POTENTIAL).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M19143; AAA29653.1; -.
CC PIR; A54498; A54498.
CC InterPro; IPR000561; EGF-like.
CC Pfam; PF00008; EGF; 1.
CC Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
CC Transmembrane; GPI-anchor.
CC SIGNAL 1 19 POTENTIAL.
CC CHAIN 20 1701 MEROZOITE SURFACE PROTEIN 1.
FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 470 470 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 802 802 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 899 899 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 919 919 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 965 965 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 991 991 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1089 1089 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT CARBOHYD 1196 1196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1588 1588 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1701 AA; 193719 MW; 3920B75E73D38552 CRC64;

Query Match 100.0%; Score 88; DB 1; Length 1701;
Best Local Similarity 100.0%; Pred. No. 6.9e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYVEKVLAKYKDDLE 17
| | | | | | | | | | | | | | | | | |
DB 1425 GISYVEKVLAKYKDDLE 1441

RESULT 3
MSP1_PLAFF STANDARD; PRT; 1701 AA.
AC P08569;
DT 01-AUG-1988 (Rel. 08, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
DE (PMMSA) (P190).
GN MSP-1.
OS Plasmodium falciparum (isolate mad20 / Papua New Guinea).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=70153;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88011243; PubMed=3079521;
RA Tanabe K., Mackay M., Goman M., Scaife J.G.;
RT "Allelic dimorphism in a surface antigen gene of the malaria parasite
RT Plasmodium falciparum."
RL J. Mol. Biol. 195:273-287(1987).
RN [2]
RP REVISIONS TO 1403; 1569 AND 1629.
RA Tanabe K.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-115 FROM N.A.
RX MEDLINE=86136024; PubMed=3004972;
RA Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,
RA Stunnenberg H., Bujard H.;
RT "Polymorphism of the precursor for the major surface antigens of
RT Plasmodium falciparum merozoites: studies at the genetic level."
RL EMBO J. 4:3823-3829(1985).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
(POTENTIAL).
CC -1- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42
CC KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
CC MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X05624; CAA29112.1; -.
CC PIR; A26868; A26868.
CC PIR; B25120; B25120.
CC InterPro; IPR000561; EGF-like.
CC Pfam; PF00008; EGF; 1.
CC Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
CC Transmembrane; GPI-anchor.
CC SIGNAL 1 19 POTENTIAL.
CC CHAIN 20 1701 MEROZOITE SURFACE PROTEIN 1.
FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 470 470 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

FT CARBOHYD 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 802 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 899 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 919 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 965 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 991 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1089 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1588 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1701 AA; 193768 MW; 3FC2EC59AF96EA98 CRC64;

Query Match 100.0%; Score 88; DB 1; Length 1701;
Best Local Similarity 100.0%; Pred. No. 6.9e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKEKVLAKYKDDLE 17
IIIIIIIIIIIIIIIIIIII

Db 1425 GISYKEKVLAKYKDDLE 1441

RESULT 4
MSPL_PLAFC
ID MSPL_PLAFC STANDARD; PRT; 1726 AA.
AC P04934;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
DE (PMMSA) (P195).
GN MSP-1.
OS Plasmodium falciparum (isolate Camp / Malaysia).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5835;
RN [1]
RP SEQUENCE OF 1-1103 FROM N.A.
RX MEDLINE=86205236; PubMed=3517809;
RA Weber J.L., Leininger W.M., Lyon J.A.;
RT "Variation in the gene encoding a major merozoite surface antigen of the human malaria parasite Plasmodium falciparum.";
RL Nucleic Acids Res. 14:3311-3323(1986).
RN [2]
RP SEQUENCE OF 1104-1726 FROM N.A.
RX MEDLINE=88143999; PubMed=3278296;
RA Weber J.L., Slim B.K.L., Lyon J.A., Wolff R.;
RT "Merozoite surface protein sequence from the Camp strain of the human malaria parasite Plasmodium falciparum.";
RL Nucleic Acids Res. 16:1206-1206(1988).
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (POTENTIAL).
CC -!- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X03831; CAA27446.1; -
DR PIR; A23386; SAZOGM.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchor.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1726 MEROZOITE SURFACE PROTEIN 1.
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 567 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 638 638 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 827 827 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 924 924 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 944 944 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 990 990 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1016 1016 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1114 1114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1221 1221 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1613 1613 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1726 AA; 196197 MW; DD8AD45FA352BCF3 CRC64;

Query Match 100.0%; Score 88; DB 1; Length 1736;
Best Local Similarity 100.0%; Pred. No. 7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKEKVLAKYKDDLE 17
IIIIIIIIIIIIIIIIIIII

Db 1450 GISYKEKVLAKYKDDLE 1466

RESULT 5
MSPL_PLAFC
ID MSPL_PLAFC STANDARD; PRT; 1726 AA.
AC P50495;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
DE (PMMSA) (GP195).
GN MSP-1.
OS Plasmodium falciparum (isolate Palo Alto / Uganda).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57270;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89005525; PubMed=3049134;
RA Chang S.P., Kramer K.J., Yamaga K.M., Kato A., Case S.E.,
RA Siddiqui W.A.;
RT "Plasmodium falciparum: gene structure and hydropathy profile of the major merozoite surface antigen (gp195) of the Uganda-Palo Alto isolate.";
RL Exp. Parasitol. 67:1-11(1988).
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (POTENTIAL).
CC -!- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M37213; AAA29611.1; -
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchor.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1726 MEROZOITE SURFACE PROTEIN 1.
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 638 638 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 827 827 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 924 924 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 944 944 N-LINKED (GLCNAC. . .) (POTENTIAL).

```
FT CARBOHYD 990 990 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1016 1016 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1114 1114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1221 1221 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1613 1613 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1726 AA; 196174 MW; 5859CEEFA2F9A026 CRC64;

Query Match 100.0%; Score 88; DB 1; Length 1726;
Best Local Similarity 100.0%; Pred. No. 7e-06; 0; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 0;

QY 1 GISYKVLAKYKDDLE 17
| | | | | | | | | | | | | | | | | |
Db 1450 GISYKVLAKYKDDLE 1466

RESULT 6
HISX_LACLA STANDARD; PRT; 431 AA.
ID HISX_LACLA
AC Q02136;
DT 01-JUL-1993 (Rel. 26, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HISTIDINOL DEHYDROGENASE (EC 1.1.1.23) (HDH).
GN HISD.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCDO 2118;
RX MEDLINE=93015709; PubMed=1400209;
RA DeLorme C., Ehrlich S.D., Renault P.;
RT "Histidine biosynthesis genes in Lactococcus lactis subsp. lactis.";
RL J. Bacteriol. 174:6571-6579(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
CC -!- FUNCTION: THIS PROTEIN IS CONSIDERED AS A BIFUNCTIONAL ENZYME,
CC POSSESSING TWO ACTIVE SITES, ONE AN ALCOHOL DEHYDROGENASE AND
CC THE OTHER AN ALDEHYDE DEHYDROGENASE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: L-HISTIDINOL + 2 NAD(+) = L-HISTIDINE +
CC 2 NADH.
CC -!- PATHWAY: TENTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SIMILARITY: TO OTHER PROKARYOTIC, FUNGAL AND PLANTS HDH.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U92974; ; NOT ANNOTATED_CDS.
CC EMBL; A006353; AAK05307.1; .
CC PIR; E45734; E45734.
CC InterPro; IPR001692; Histidinol_dh.
CC Pfam; PF00815; Histidinol_dh; 1.
CC PRINTS; PR00083; HOLDHIDRGNASE.
CC ProDom; PD002680; Histidinol_dh; 1.
CC PROSITE; PS00611; HISOL_DEHYDROGENASE; 1.
CC Histidine biosynthesis; Multifunctional enzyme; Oxidoreductase; NAD;
CC Complete proteome.
KW
```

```
FT CONFLICT 25 25 L -> V (IN REF. 1).
FT CONFLICT 38 38 E -> A (IN REF. 1).
FT CONFLICT 43 43 S -> N (IN REF. 1).
FT CONFLICT 69 69 P -> R (IN REF. 1).
FT CONFLICT 74 74 T -> A (IN REF. 1).
FT CONFLICT 164 164 T -> A (IN REF. 1).
FT CONFLICT 245 245 E -> K (IN REF. 1).
FT CONFLICT 248 248 N -> K (IN REF. 1).
FT CONFLICT 295 295 S -> R (IN REF. 1).
FT CONFLICT 301 301 Y -> S (IN REF. 1).
FT CONFLICT 303 303 S -> R (IN REF. 1).
FT CONFLICT 316 316 E -> D (IN REF. 1).
FT CONFLICT 403 403 A -> D (IN REF. 1).
SQ SEQUENCE 431 AA; 47352 MW; 50D7E486CBF3613C CRC64;

Query Match 50.6%; Score 44.5; DB 1; Length 431;
Best Local Similarity 62.5%; Pred. No. 11;
Matches 10; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 3 SYVEK-VLAKYKDDLE 17
| | | | | | | | | | | | | | | | | |
Db 394 SYYPEKVLAKYKDDLE 409

RESULT 7
Y069_CHLTR STANDARD; PRT; 451 AA.
ID Y069_CHLTR
AC 084072;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE METAL TRANSPORT SYSTEM MEMBRANE PROTEIN CT069.
GN CT069.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/CX;
RX MEDLINE=9900809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -!- FUNCTION: PART OF AN ATP-DRIVEN TRANSPORT SYSTEM
CC CT067/CT068/CT069/CT070 FOR A METAL.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (PROBABLE).
CC -!- SIMILARITY: BELONGS TO THE ABC-3 SUBFAMILY OF INTEGRAL MEMBRANE
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001281; AAC67660.1; .
CC InterPro; IPR001367; HTH_DtXR.
CC DR InterPro; IPR001626; ABC-3.
CC DR Pfam; PF00950; ABC-3; 1.
CC DR Pfam; PF01325; Fe-dep.repress; 1.
CC DR SMART; SM00529; HTH_DtXR; 1.
CC KW Hypothetical protein; Transport; Transmembrane; Inner membrane;
CC Complete proteome.
CC FT TRANSMEM 14 34 POTENTIAL.
CC FT TRANSMEM 38 58 POTENTIAL.
CC FT TRANSMEM 70 90 POTENTIAL.
```


FT TRANSMEM 100 120 POTENTIAL.
FT TRANSMEM 145 165 POTENTIAL.
FT TRANSMEM 192 212 POTENTIAL.
FT TRANSMEM 233 253 POTENTIAL.
FT TRANSMEM 269 289 POTENTIAL.
SQ SEQUENCE 451 AA; 51231 MW; 925F4BD18C473C5B CRC64;

Query Match 50.0%; Score 44; DB 1; Length 451;
Best Local Similarity 64.3%; Pred. No. 14;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKD 14
Db 86 GISFLEKLAHMD 99
|||||

RESULT 8
YID6_YEAST STANDARD; PRT; 587 AA.
AC P40535;

DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE HYPOTHETICAL 65.3 KDA PROTEIN IN NOT3-CRA1 INTERGENIC REGION.
GN YIL036W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=S288C / Ab972;
RA Barrall B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO YEAST YER045C.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; 246861; CAAB6915.1; -;
DR SGD; S0001298; CST6.
DR Pfam; PF00170; BZIP; 1.
DR SMART; SM00338; BRLZ; 1.
KW Hypothetical protein; ATP-binding.
FT NP_BIND 382 389 ATP (POTENTIAL).
SQ SEQUENCE 587 AA; 65264 MW; 1B98DC38BC8CAE94 CRC64;

Query Match 50.0%; Score 44; DB 1; Length 587;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 2 ISYKVLAKYK 13
Db 471 LNYEKLISKFK 482
|||||

RESULT 9

MURD_RICPR STANDARD; PRT; 445 AA.
AC Q92DC2;
DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE UDP-N-ACETYLURAMUOYLALANINE--D-GLUTAMATE LIGASE (EC 6.3.2.9) (UDP-N-
DE ACETYLURAMUOYL-L-ALANYL-D-GLUTAMATE SYNTHETASE) (D-GLUTAMIC ACID
DE ADDING ENZYME).
GN MURD OR RP410.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sichertitz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria";
RL Nature 396:133-140(1998).
CC -!- FUNCTION: CELL WALL FORMATION. CATALYZES THE ADDITION OF D-
CC GLUTAMATE TO THE NUCLEOTIDE PRECURSOR UDP-N-ACETYLURAMUOYL-L-
CC ALANINE (UMA) (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + UDP-N-ACETYLURAMUOYL-L-ALANINE +
CC D-GLUTAMATE = ADP + ORTHOPHOSPHATE + UDP-N-ACETYLURAMUOYL-
CC L-ALANYL-D-GLUTAMATE.
CC -!- PATHWAY: PEPTIDOGLYCAN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE MURDEF FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AJ235271; CAA14867.1; -;
DR InterPro; IPR000713; Mur_Ligase.
DR Pfam; PF01225; Mur_Ligase; 1.
KW Peptidoglycan synthetase; Cell wall; Cell division; Ligase;
KW ATP-binding; Complete proteome.
FT NP_BIND 111 117 ATP (POTENTIAL).
SQ SEQUENCE 445 AA; 50129 MW; B9CCCF7437FB7AA6 CRC64;

Query Match 48.9%; Score 43; DB 1; Length 445;
Best Local Similarity 52.4%; Pred. No. 19;
Matches 11; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

QY 1 GISYKVLAKYK---DDLE 17
Db 18 GISVVEELKNKYDLIVYDDLE 38
|||||

RESULT 10

YMP9_CAEEL STANDARD; PRT; 194 AA.
ID YMP9_CAEEL
AC Q10952;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 22.0 KDA PROTEIN B0361.9 IN CHROMOSOME III PRECURSOR.
GN B0361.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Du Z.;

RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U00031; AAK18868.1; -
DR WormPep; B0361.9; CE00838.
DR InterPro; IPR003014; PAN.
DR Pfam; PF00024; PAN; 1.
DR SMART; SM00473; PAN.AP; 1.
DR Hypothetical protein; Signal.
KW SIGNAL
FT CHAIN 1 15 POTENTIAL.
FT CHAIN 16 194 HYPOTHETICAL PROTEIN B0361.9.
SQ SEQUENCE 194 AA; 22017 MW; 79C9F7A2570AAF38 CRC64;

Query Match 47.7%; Score 42; DB 1; Length 194;
Best Local Similarity 46.7%; Pred. No. 12;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 SYIEKVLAKYKDDLE 17
DB 127 TYIEKIRCLDDVE 141
:||||: :||:|

RESULT 11
CAB1_CAEEL STANDARD; PRT; 425 AA.
ID CAB1_CAEEL
AC Q93249; P90751;
DT 15-JUL-1998 (Rel. 36, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROTEIN CAB-1.
DE CAB-1 OR C23H4.1.
GN CAB-1 OR C23H4.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderiinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A., AND INTERACTION WITH AEX-3.
RC STRAIN-BRISTOL N2;
RX MEDLINE=20428446; PubMed=10970871;
RA Iwasaki K., Toyonaga R.;
RT "The rab3 GDP/GTP exchange factor homolog AEX-3 has a dual function
in synaptic transmission."
RL EMBO J. 19:4806-4816(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Wilkinson J.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS
RC STRAIN-BRISTOL N2;
RA Durbin R.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: BINDS TO THE RAB3 GDP/GTP EXCHANGE FACTOR AEX-3.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF NEURONS.
CC -1- SIMILARITY: BELONGS TO THE NPDC1 / CAB-1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----

CC EMBL; AF293976; AAG17881.1; -
DR EMBL; Z784116; CAB01682.1; -
DR WormPep; C23H4.1; CE17435.
KW Transmembrane.
FT TRANSMEM 300 320 POTENTIAL.
FT DOMAIN 205 424 AEX-3-BINDING.
SQ SEQUENCE 425 AA; 47224 MW; 6CC2C50AEEB7A1FA CRC64;

Query Match 47.7%; Score 42; DB 1; Length 425;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISYIEKVLAKYKDD 15
DB 97 LKYYGKEVEQYKDD 110
:||||: :||||

RESULT 12
THII_ECOLI STANDARD; PRT; 482 AA.
ID THII_ECOLI
AC P77718;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE THIAMINE BIOSYNTHESIS PROTEIN THII.
GN THII OR B0423.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,
RA Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
RA Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 253-482 FROM N.A.
RC STRAIN-K12;
RA Backstrom A.D.;
RL Thesis (1996), Cornell University, U.S.A.
CC -1- FUNCTION: REQUIRED FOR THE SYNTHESIS OF THE THIAZOLE MOIETY.
CC -1- PATHWAY: THIAMINE BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE THII FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000148; AAC73526.1; -
DR EMBL; U82664; AAB40179.1; -
DR EMBL; U34923; -; NOT_ANNOTATED_CDS.
DR EcoGene; EG13273; thii.
DR InterPro; IPR003720; Thii.
DR Pfam; PF02568; Thii; 1.
DR

```
KW Thiamine biosynthesis; Complete proteome.
SQ SEQUENCE 482 AA; 54973 MW; 72AF449459E45762 CRC64;

Query Match 47.7%; Score 42; DB 1; Length 482;
Best Local Similarity 53.8%; Pred. No. 30;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 YEKVLAKYKDDLE 17
DB 91 FEKALVQYRDQLE 103

RESULT 13
THII_SALTY STANDARD; PRT; 482 AA.
AC P55913; O06955;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE THIAMINE BIOSYNTHESIS PROTEIN THII.
GN THII.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=97352701; PubMed=9209060;
RA Webb E., Claas K., Downs D.M.;
RT "Characterization of thii, a new gene involved in thiazole
biosynthesis in Salmonella typhimurium.";
RL J. Bacteriol. 179:4399-4402(1997).
RN [2]
RP SEQUENCE OF 365-482 FROM N.A.
RA Metcalf W.W., Jiang W., Wanner B.L.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: REQUIRED FOR THE SYNTHESIS OF THE THIAZOLE MOIETY.
CC -!- PATHWAY: THIAMINE BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE THII FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; U69493; AAB39648.1; -
DR EMBL; U94901; AAB63031.1; -
DR Stygene; SG10723; thii.
DR InterPro; IPR003720; Thii.
DR Pfam; PF02568; Thii; 1.
KW Thiamine biosynthesis.
FT CONFLICT 365 365 E -> R (IN REF. 2).
SQ SEQUENCE 482 AA; 54831 MW; 41CC9E3740CE5D2 CRC64;

Query Match 47.7%; Score 42; DB 1; Length 482;
Best Local Similarity 53.8%; Pred. No. 30;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 YEKVLAKYKDDLE 17
DB 91 FEKALVQYRDQLE 103

RESULT 14
VMCB_LAMBD STANDARD; PRT; 533 AA.
ID VMCB_LAMBD
```

```
AC P03710;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
DE PORTAL PROTEIN (GPB) (MINOR CAPSID PROTEIN B).
GN B.
OS Bacteriophage lambda.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;
OC Lambda phage group.
OX NCBI_TaxID=10710;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83189071; PubMed=6221115;
RA Sanger F., Coulson A.R., Hong G.F., Hill D.F., Petersen G.B.;
RT "Nucleotide sequence of bacteriophage lambda DNA.";
RL J. Mol. Biol. 162:729-773(1982).
CC -!- FUNCTION: FORMS THE PORTAL VERTEX OF THE PROHEAD, THE CENTRAL
HOLE OF THE PORTAL VERTEX IS THE OPENING THROUGH WHICH DNA PASSES
DURING PACKAGING AND INJECTION.
CC -!- SUBUNIT: DODECAMER OF PROTEIN B AND PROTEIN B*.
CC -!- PTM: PROTEIN B* IS A CLEAVAGE PRODUCT OF PROTEIN B.
CC -!- SIMILARITY: TO BACTERIOPHAGE 21 HEAD PROTEIN GP4.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-15 IS THE INITIATOR.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; J02459; AAA96536.1; -
DR PIR; A04332; VHBPBL.
KW Coat protein; DNA-binding.
FT CHAIN 1 533 PROTEIN B.
FT CHAIN 23 533 PROTEIN B*.
FT DNA_BIND 359 378 H-T-H MOTIF (POTENTIAL).
SQ SEQUENCE 533 AA; 59465 MW; 4E8E25B30A9144A0 CRC64;

Query Match 47.7%; Score 42; DB 1; Length 533;
Best Local Similarity 52.9%; Pred. No. 34;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 GISYVEKVLAKYKDDLE 17
DB 471 GLSTVEKECAKRGDDYQ 487

RESULT 15
NIAL_BRANA STANDARD; PRT; 911 AA.
ID NIAL_BRANA
AC P39867;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE NITRATE REDUCTASE, CLONE PNBRI405 (EC 1.6.6.1) (NR).
GN NIAL.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LISANDRA;
RX MEDLINE=96222419; PubMed=8685274;
RA Fukuoka H., Ogawa T., Minami H., Yano H., Ohkawa Y.;
RT "Developmental stage-specific and nitrate-independent regulation of
nitrate reductase gene expression in rapeseed.";
RL Plant Physiol. 111:39-47(1996).
CC -!- FUNCTION: NITRATE REDUCTASE IS A KEY ENZYME INVOLVED IN THE FIRST
```

```

CC STEP OF NITRATE ASSIMILATION IN PLANTS, FUNGI AND BACTERIA.
CC -|- CATALYTIC ACTIVITY: NADH + NITRATE = NAD(+) + NITRITE + H(2)O.
CC -|- COFACTOR: REQUIRES FAD, A HEME GROUP (CALLED CYTOCHROME B-557)
CC AND ONE MOLYBDENUM ATOM.
CC -|- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -|- SIMILARITY: TO EUKARYOTIC MOLYBDOPTERIN OXIDOREDUCTASES IN THE
CC N-TERMINAL DOMAIN.
CC -|- SIMILARITY: CONTAINS A CYTOCHROME B5 FAMILY, HEME-BINDING DOMAIN.
CC -|- SIMILARITY: TO FAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE
CC C-TERMINAL DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D38219; BAA07394.1; -.
CC HSSP: P17571; 2CND.
CC Mendel; 7474; BRANA:Nial;1.
CC InterPro: IPR001199; Cyt_B5.
CC InterPro: IPR001834; Cyt_B5_reductase.
CC InterPro: IPR000572; Euk_oxidored_molzb.
CC InterPro: IPR001709; Flavpyrid_cyt_redctse.
CC InterPro: IPR001433; Oxidored_FAD.
CC Pfam; PF00970; Cyt_reductase; 1.
CC Pfam; PF00173; heme_1; 1.
CC Pfam; PF00175; oxidored_fad; 1.
CC Pfam; PF00174; oxidored_molzb; 1.
CC PRINTS: PR00363; CYTOCHROME_B5.
CC PRINTS: PR00371; FENCR.
CC PRINTS: PR00406; CYTB5RDASE.
CC PRINTS: PR00407; EUOPTERIN.
CC PROSITE; PS00191; CYTOCHROME_B5_1; 1.
CC PROSITE; PS50255; CYTOCHROME_B5_2; 1.
CC PROSITE; PS00559; MOLYBDOPTERIN_EUK; 1.
CC Oxidoreductase; Flavoprotein; FAD; NAD; Heme; Molybdenum;
CC Nitrate assimilation; Multigene family.
CC METAL. 191 191 MOLYBDENUM-PTERIN (POTENTIAL).
CC METAL 245 245 MOLYBDENUM-PTERIN (POTENTIAL).
CC DISULFID 430 430 INTERCHAIN (POTENTIAL).
CC BINDING 574 574 HEME LIGAND (BY SIMILARITY).
CC BINDING 597 597 HEME LIGAND (BY SIMILARITY).
CC SEQUENCE 911 AA; 102252 MW; BBA4E19835B03D8C CRC64;

```

```

Query Match 47.7%; Score 42; DB 1; Length 911;
Best Local Similarity 43.8%; Pred. No. 58;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISYKVLAKYKDDLE 17
Db :|||||:|
74 VSYKENVLKNSDLE 89

```

Search completed: January 29, 2002, 11:13:44
Job time: 820 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:12:12 ; Search time 285.36 seconds
(without alignments)
8.714 Million cell updates/sec

Title: US-09-763-397A-13
Perfect score: 88
Sequence: 1 GISYKEVLAKYKDDLE 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	88	100.0	372	5	Q43997
2	88	100.0	372	5	Q25717
3	88	100.0	372	5	Q25718
4	88	100.0	372	5	Q25725
5	88	100.0	372	5	Q25726
6	88	100.0	372	5	Q25719
7	88	100.0	372	5	Q25720
8	88	100.0	373	5	Q43995
9	88	100.0	373	5	Q43996
10	88	100.0	373	5	Q25721
11	88	100.0	373	5	Q25722
12	88	100.0	373	5	Q25723
13	88	100.0	373	5	Q25724
14	88	100.0	373	5	Q25727
15	88	100.0	373	5	Q25728
16	88	100.0	569	5	Q25967
17	88	100.0	569	5	Q25969
18	88	100.0	569	5	Q25970
19	88	100.0	569	5	Q25974

20	88	100.0	569	5	Q25975	Q25975 plasmodium
21	88	100.0	569	5	Q25977	Q25977 plasmodium
22	88	100.0	569	5	Q25978	Q25978 plasmodium
23	88	100.0	569	5	Q25979	Q25979 plasmodium
24	88	100.0	569	5	Q25980	Q25980 plasmodium
25	88	100.0	569	5	Q25982	Q25982 plasmodium
26	88	100.0	569	5	Q25983	Q25983 plasmodium
27	88	100.0	570	5	Q25968	Q25968 plasmodium
28	88	100.0	570	5	Q9TYG2	Q9TYG2 plasmodium
29	88	100.0	651	5	Q25924	Q25924 plasmodium
30	88	100.0	652	5	Q25923	Q25923 plasmodium
31	88	100.0	1694	5	Q9TZT5	Q9TZT5 plasmodium
32	88	100.0	1694	5	Q9NHX1	Q9NHX1 plasmodium
33	88	100.0	1704	5	Q9TZT4	Q9TZT4 plasmodium
34	88	100.0	1720	5	Q25922	Q25922 plasmodium
35	50	56.8	815	10	Q9FZ65	Q9FZ65 arabidopsis
36	47	53.4	806	5	P91136	P91136 caenorhabdi
37	46	52.3	404	2	Q9AGR4	Q9AGR4 streptococc
38	45	51.1	1011	4	Q9UP95	Q9UP95 homo sapien
39	45	51.1	1068	4	O60632	O60632 homo sapien
40	45	51.1	1085	4	Q13953	Q13953 homo sapien
41	45	51.1	2470	5	O9VK45	O9VK45 drosophila
42	44	50.0	136	2	Q56526	Q56526 unidentified
43	44	50.0	136	2	Q56528	Q56528 unidentified
44	44	50.0	265	2	Q9SIF1	Q9SIF1 vibrio chol
45	44	50.0	325	5	Q9BID7	Q9BID7 branchiosto

ALIGNMENTS

RESULT 1	
O43997	
ID O43997	PRELIMINARY; PRT; 372 AA.
AC	
DT 01-JUN-1998	(TREMBLrel. 06, Created)
DT 01-JUN-1998	(TREMBLrel. 06, Last sequence update)
DT 01-JUN-2001	(TREMBLrel. 17, Last annotation update)
DE	MEROZOITE SURFACE PROTEIN-1 (FRAGMENT).
GN MSP-1.	
OS	Plasmodium falciparum.
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX	NCBI_TaxID=5833;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=KENYA-3;
RA	Qari S.H., Shi Y.P., Goldman I., Nahlen B., Tibayrenc M., Lal A.A.;
RL	Mol. Biochem. Parasitol. 0:0-0(1998).
DR	EMBL; AF040569; AAC39099.1; -
DR	InterPro; IPR000561; EGF-like.
DR	Pfam; PF00008; EGF; 1.
KW	Merozoite.
FT	NON_TER 1
FT	NON_TER 372
SQ	SEQUENCE 372 AA; 42768 MW; 3ACD3AFA6C047D53 CRC64;
Query Match	100.0%; Score 88; DB 5; Length 372;
Best Local Similarity	100.0%; Pred. No. 3.le-05;
Matches 17; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY 1	GISYKEVLAKYKDDLE 17
Db 118	GISYKEVLAKYKDDLE 134
RESULT 2	
ID Q25717	PRELIMINARY; PRT; 372 AA.
AC	
DT 01-NOV-1996	(TREMBLrel. 01, Created)
DT 01-NOV-1996	(TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001	(TREMBLrel. 17, Last annotation update)

```
DE MEROZOITE SURFACE PROTEIN 1, 42 KDA C-TERMINAL REGION (FRAGMENT).
GN MSP-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN
RP SEQUENCE FROM N.A.
RC STRAIN-INFECTED HUMAN BLOOD SAMPLE FROM PAPUA NEW GUINEA;
RA Shi Y.P., Alpers M.P., Povoa M.M., Nahlen B.L., Oloo A.G., Lal A.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20653; AAA62213.1; -.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
KW Merozoite.
FT NON_TER 1 1
FT NON_TER 372 372
SQ SEQUENCE 372 AA; 42687 MW; 3F2BF1152598FB10 CRC64;

Query Match 100.0%; Score 88; DB 5; Length 372;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17
DB 118 GISYKVLAKYKDDLE 134

RESULT 3
Q25718 PRELIMINARY; PRT; 372 AA.
AC Q25718;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1, 42 KDA C-TERMINAL REGION (FRAGMENT).
GN MSP-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN
RP SEQUENCE FROM N.A.
RC STRAIN-INFECTED HUMAN BLOOD SAMPLE FROM PAPUA NEW GUINEA;
RA Shi Y.P., Alpers M.P., Povoa M.M., Nahlen B.L., Oloo A.G., Lal A.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20654; AAA62214.1; -.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
KW Merozoite.
FT NON_TER 1 1
FT NON_TER 372 372
SQ SEQUENCE 372 AA; 42660 MW; E9F088252598FB0B CRC64;

Query Match 100.0%; Score 88; DB 5; Length 372;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17
DB 118 GISYKVLAKYKDDLE 134

RESULT 4
Q25725 PRELIMINARY; PRT; 372 AA.
AC Q25725;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1, 42 KDA C-TERMINAL REGION (FRAGMENT).
GN MSP-1.
OS Plasmodium falciparum.
```

```
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN
RP SEQUENCE FROM N.A.
RC STRAIN-INFECTED HUMAN BLOOD SAMPLE FROM PAPUA NEW GUINEA;
RA Shi Y.P., Alpers M.P., Povoa M.M., Nahlen B.L., Oloo A.G., Lal A.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20730; AAA62221.1; -.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
KW Merozoite.
FT NON_TER 1 1
FT NON_TER 372 372
SQ SEQUENCE 372 AA; 42711 MW; E3DBB33F282CD5FC CRC64;

Query Match 100.0%; Score 88; DB 5; Length 372;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17
DB 118 GISYKVLAKYKDDLE 134

RESULT 5
Q25726 PRELIMINARY; PRT; 372 AA.
AC Q25726;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1, 42 KDA C-TERMINAL REGION (FRAGMENT).
GN MSP-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN
RP SEQUENCE FROM N.A.
RC STRAIN-INFECTED HUMAN BLOOD SAMPLE FROM PAPUA NEW GUINEA;
RA Shi Y.P., Alpers M.P., Povoa M.M., Nahlen B.L., Oloo A.G., Lal A.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20731; AAA62222.1; -.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
KW Merozoite.
FT NON_TER 1 1
FT NON_TER 372 372
SQ SEQUENCE 372 AA; 42734 MW; 19844F6D7011E52 CRC64;

Query Match 100.0%; Score 88; DB 5; Length 372;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17
DB 118 GISYKVLAKYKDDLE 134

RESULT 6
Q25719 PRELIMINARY; PRT; 372 AA.
AC Q25719;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1, 42 KDA C-TERMINAL REGION (FRAGMENT).
GN MSP-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN-INFECTED HUMAN BLOOD SAMPLE FROM PAPUA NEW GUINEA;
RA Shi Y.P., Alpers M.P., Povoa M.M., Nahlen B.L., Oloo A.G., Lal A.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20655; AAA62215.1; -.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
KW Merozoite.
FT NON_TER 1 1
FT NON_TER 372 372
SQ SEQUENCE 372 AA; 42686 MW; 583A33972DB05FB4 CRC64;

Query Match 100.0%; Score 88; DB 5; Length 372;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GISYKVLAKYKDDLE 17
Db 118 GISYKVLAKYKDDLE 134

RESULT 7
Q25720 ID Q25720 PRELIMINARY; PRT; 372 AA.
AC Q25720;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1, 42 KDA C-TERMINAL REGION (FRAGMENT).
GN MSP-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INFECTED HUMAN BLOOD SAMPLE FROM PAPUA NEW GUINEA;
RA Shi Y.P., Alpers M.P., Povoa M.M., Nahlen B.L., Oloo A.G., Lal A.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20656; AAA62216.1; -.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
KW Merozoite.
FT NON_TER 1 1
FT NON_TER 372 372
SQ SEQUENCE 372 AA; 42686 MW; 583A33972DB05FB4 CRC64;

Query Match 100.0%; Score 88; DB 5; Length 372;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GISYKVLAKYKDDLE 17
Db 118 GISYKVLAKYKDDLE 134

RESULT 8
Q43995 ID Q43995 PRELIMINARY; PRT; 373 AA.
AC Q43995;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-1998 (TREMBlrel. 06, Last annotation update)
DE MEROZOITE SURFACE PROTEIN-1 (FRAGMENT).
GN MSP-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KENYA-1;
RA Qari S.H., Shi Y.P., Goldman I., Nahlen B., Tibayrenc M., Lal A.A.;
```

```
RL Mol. Biochem. Parasitol. 0:0-0(1998).
DR EMBL; AF040567; AAC39097.1; -.
KW Merozoite.
FT NON_TER 1 1
FT NON_TER 373 373
SQ SEQUENCE 373 AA; 42873 MW; 886CF169A7AF5194 CRC64;

Query Match 100.0%; Score 88; DB 5; Length 373;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GISYKVLAKYKDDLE 17
Db 118 GISYKVLAKYKDDLE 134

RESULT 9
Q43996 ID Q43996 PRELIMINARY; PRT; 373 AA.
AC Q43996;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN-1 (FRAGMENT).
GN MSP-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KENYA-2;
RA Qari S.H., Shi Y.P., Goldman I., Nahlen B., Tibayrenc M., Lal A.A.;
RL Mol. Biochem. Parasitol. 0:0-0(1998).
DR EMBL; AF040568; AAC39098.1; -.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
KW Merozoite.
FT NON_TER 1 1
FT NON_TER 373 373
SQ SEQUENCE 373 AA; 42902 MW; 86CD4B721E605A5F CRC64;

Query Match 100.0%; Score 88; DB 5; Length 373;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GISYKVLAKYKDDLE 17
Db 118 GISYKVLAKYKDDLE 134

RESULT 10
Q25721 ID Q25721 PRELIMINARY; PRT; 373 AA.
AC Q25721;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1, 42 KDA C-TERMINAL REGION (FRAGMENT).
GN MSP-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INFECTED HUMAN BLOOD SAMPLE FROM PAPUA NEW GUINEA;
RA Shi Y.P., Alpers M.P., Povoa M.M., Nahlen B.L., Oloo A.G., Lal A.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20726; AAA62217.1; -.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
KW Merozoite.
```

```
FT NON_TER 1 1
FT NON_TER 373 373
SQ SEQUENCE 373 AA; 42848 MW; EE0700233D7F4D4E CRC64;

Query Match 100.0%; Score 88; DB 5; Length 373;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17
Db 118 GISYKVLAKYKDDLE 134

RESULT 11
Q25722 ID Q25722 PRELIMINARY; PRT; 373 AA.
AC Q25722;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE DE MEROZOITE SURFACE PROTEIN 1, 42 KDA C-TERMINAL REGION (FRAGMENT).
GN MSP-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INFECTED HUMAN BLOOD SAMPLE FROM PAPUA NEW GUINEA;
RA Shi Y.P., Alpers M.P., Povoa M.M., Nahlen B.L., Oloo A.G., Lal A.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20727; AAA62218.1; -.
DR InterPro: IPR000561; EGF-like.
DR Pfam: PF00008; EGF; 1.
KW Merozoite.
FT NON_TER 1 1
FT NON_TER 373 373
SQ SEQUENCE 373 AA; 42871 MW; D29DC2517DDE1B4A CRC64;

Query Match 100.0%; Score 88; DB 5; Length 373;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17
Db 118 GISYKVLAKYKDDLE 134

RESULT 12
Q25723 ID Q25723 PRELIMINARY; PRT; 373 AA.
AC Q25723;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE DE MEROZOITE SURFACE PROTEIN 1, 42 KDA C-TERMINAL REGION (FRAGMENT).
GN MSP-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INFECTED HUMAN BLOOD SAMPLE FROM PAPUA NEW GUINEA;
RA Shi Y.P., Alpers M.P., Povoa M.M., Nahlen B.L., Oloo A.G., Lal A.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20728; AAA62219.1; -.
DR InterPro: IPR000561; EGF-like.
DR Pfam: PF00008; EGF; 1.
KW Merozoite.
FT NON_TER 1 1
FT NON_TER 373 373
SQ SEQUENCE 373 AA; 42848 MW; EE9A891631DE174F CRC64;

Query Match 100.0%; Score 88; DB 5; Length 373;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17
Db 118 GISYKVLAKYKDDLE 134

RESULT 13
Q25724 ID Q25724 PRELIMINARY; PRT; 373 AA.
AC Q25724;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE DE MEROZOITE SURFACE PROTEIN 1, 42 KDA C-TERMINAL REGION (FRAGMENT).
GN MSP-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INFECTED HUMAN BLOOD SAMPLE FROM PAPUA NEW GUINEA;
RA Shi Y.P., Alpers M.P., Povoa M.M., Nahlen B.L., Oloo A.G., Lal A.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20729; AAA62220.1; -.
DR InterPro: IPR000561; EGF-like.
DR Pfam: PF00008; EGF; 1.
KW Merozoite.
FT NON_TER 1 1
FT NON_TER 373 373
SQ SEQUENCE 373 AA; 42839 MW; A21A91B2076F415A CRC64;

Query Match 100.0%; Score 88; DB 5; Length 373;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17
Db 118 GISYKVLAKYKDDLE 134

RESULT 14
Q25727 ID Q25727 PRELIMINARY; PRT; 373 AA.
AC Q25727;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE DE MEROZOITE SURFACE PROTEIN 1, 42 KDA C-TERMINAL REGION (FRAGMENT).
GN MSP-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INFECTED HUMAN BLOOD SAMPLE FROM PAPUA NEW GUINEA;
RA Shi Y.P., Alpers M.P., Povoa M.M., Nahlen B.L., Oloo A.G., Lal A.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20732; AAA62223.1; -.
DR InterPro: IPR000561; EGF-like.
DR Pfam: PF00008; EGF; 1.
KW Merozoite.
FT NON_TER 1 1
FT NON_TER 373 373
SQ SEQUENCE 373 AA; 42814 MW; A799666CF38CC405C CRC64;

Query Match 100.0%; Score 88; DB 5; Length 373;
```


Best Local Similarity 100.0%; Pred. No. 3.le-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKVKLAKYKDDLE 17
DB 118 GISYKVKLAKYKDDLE 134

RESULT 15
Q25728 PRELIMINARY; PRT; 373 AA.
AC Q25728;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1, 42 KDA C-TERMINAL REGION (FRAGMENT).
GN MSP-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INFECTED HUMAN BLOOD SAMPLE FROM PAPUA NEW GUINEA;
RA Shi Y.P., Alpers M.P., Pova M.M., Nahlen B.L., Oloo A.G., Lal A.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20733; AAA62224.1;
DR InterPro: IPR000561; EGF-like.
DR Pfam: PF00008; EGF; 1.
KW Merozoite.
FT NON_TER 1
FT NON_TER 373
SQ SEQUENCE 373 AA; 42815 MW; 9CF4DDAF38CC4054 CRC64;

Query Match 100.0%; Score 88; DB 5; Length 373;
Best Local Similarity 100.0%; Pred. No. 3.le-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKVKLAKYKDDLE 17
DB 118 GISYKVKLAKYKDDLE 134

Search completed: January 29, 2002, 11:12:13
Job time: 769 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:21:46 ; Search time 310.82 Seconds
(without alignments)
1.907 Million cell updates/sec

Title: US-09-763-397A-14
Perfect score: 41
Sequence: 1 SNTFINNA 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	8	21	AAV70290
2	41	100.0	264	9	AAV80365
3	41	100.0	272	11	AAV05876
4	41	100.0	272	21	AAV18202
5	41	100.0	280	11	AAV05879
6	41	100.0	287	11	AAV05877
7	41	100.0	300	11	AAV05878
8	41	100.0	350	21	AAV070278
9	37	90.2	17	11	AAV03985
10	34	82.9	469	11	AAV04944
11	34	82.9	470	8	AAV07012

12	33	80.5	2010	21	AAV18218	Plasmodium falcipa
13	32	78.0	68	15	AAV46082	DEAD ATP helicase
14	32	78.0	216	22	AAV98434	Papio hamadryas ol
15	32	78.0	216	22	AAV98448	Gorilla olfactory
16	32	78.0	216	22	AAV72825	Human olfactory re
17	32	78.0	319	22	AAV72346	Human OR-like poly
18	32	78.0	320	22	AAV72159	Human olfactory re
19	32	78.0	320	22	AAV72168	Human olfactory re
20	32	78.0	324	22	AAV17324	Peptide #3758 enco
21	32	78.0	324	22	AAV29823	Peptide #3860 enco
22	32	78.0	324	22	AAV05005	Peptide #3687 enco
23	32	78.0	473	15	AAV52604	Allinase encoded b
24	32	78.0	481	21	AAV03415	Corn putative carb
25	32	78.0	614	22	AAV70766	S cerevisiae apopt
26	32	78.0	641	22	AAE00815	Aspergillus oryzae
27	32	78.0	678	22	AAV73379	Moraxella catarrha
28	31	75.6	143	21	AAV33925	Arabidopsis thalia
29	31	75.6	152	21	AAV33924	Arabidopsis thalia
30	31	75.6	176	10	AAV90414	Plasmodium falcipa
31	31	75.6	190	21	AAV33923	Plasmodium falcipa
32	31	75.6	215	22	AAV98475	Salmiri boliviensi
33	31	75.6	215	22	AAV98479	Salmiri sciureus o
34	31	75.6	215	22	AAV98481	Salmiri sciureus o
35	31	75.6	216	22	AAV98466	Callithrix jacchus
36	31	75.6	216	22	AAV98595	Mouse olfactory re
37	31	75.6	275	21	AAV16532	Arabidopsis thalia
38	31	75.6	276	22	AAV72193	Human olfactory re
39	31	75.6	328	21	AAV07255	Arabidopsis thalia
40	31	75.6	337	21	AAV07254	Arabidopsis thalia
41	31	75.6	366	21	AAV16531	Arabidopsis thalia
42	31	75.6	374	21	AAV16530	Arabidopsis thalia
43	31	75.6	379	21	AAV07253	Arabidopsis thalia
44	31	75.6	499	19	AAV76308	Polyporus pmsitus
45	31	75.6	499	19	AAV60876	Polyporus pmsitus

ALIGNMENTS

RESULT 1

AAV70290
ID AAV70290 standard; peptide; 8 AA.

AC AAV70290;

DT 06-JUN-2000 (first entry)

DE Plasmodium falciparum MSP-2 antigenic epitope, P543.

Recombinant protein; CDC/NIAIDVAC-1; multivalent; malaria; vaccine;
T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
Pf27; antiparasitic; prevention; anti-CDC/NIAIDVAC-1 antibody.

OS Plasmodium falciparum.

PN WO200011179-A1.

PD 02-MAR-2000.

PF 19-AUG-1999; 99WO-US18869.

PR 21-AUG-1998; 98US-0097703.

PA (NAIM-) NAT INST IMMUNOLOGY.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Lal AA, Shi YP, Hasnain SE;

XX WPI; 2000-237654/20.

XX Novel recombinant protein as vaccine for treating malarial infection
 PT comprises antigenic peptides obtained from different stages of
 PT Plasmodium falciparum life cycle -
 XX Claim 2; Page 17; 52pp; English.
 PS
 XX The present sequence is the antigenic epitope P543, derived from
 CC merozoite surface protein-2 (MSP-2) of the asexual blood stage of
 CC Plasmodium falciparum. It is used in the construction of recombinant
 CC protein CDC/NIIMALVAC-1, which is a multivalent, multistage malarial
 CC vaccine. The recombinant protein comprises, melittin signal peptide,
 CC (His)6 tag, T-cell epitope from tetanus toxoid and 21 antigenic epitopes
 CC from circumsporozoite protein (CSP), sporozoite surface protein-2
 CC (SSP-2), liver stage antigen-1 (LSA-1), merozoite surface protein-1
 CC (MSP-1), MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding
 CC antigen-175 (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete
 CC specific antigen, Pfg27. These epitopes were obtained at different stages
 CC of the life cycle of P. falciparum. CDC/NIIMALVAC-1 vaccine has
 CC antiparasitic activity and can be used for treatment and prevention of
 CC malarial infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for
 CC detecting P. falciparum in biological samples.
 XX Sequence 8 AA;
 SQ

Query Match 100.0%; Score 41; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 4.3e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNTFINNA 8
 Db 1 sntfinna 8

RESULT 2
 AAP80365
 ID AAP80365 standard; protein; 264 AA.
 XX
 AC AAP80365;
 XX
 DT 10-DEC-1990 (first entry)
 XX
 XX Sequence of Ag513 a putative merozoite surface antigen of
 DE Plasmodium falciparum.
 DE
 XX Malaria therapy; diagnosis.
 KW
 XX Plasmodium falciparum.
 OS
 XX Key Location/Qualifiers
 FH Peptide 6..20
 FT /note="Core of a signal sequence"
 FT Region 60..123
 FT /note="It contains two identical copies of a 32 AA unit
 FT arranged in tandem"
 FT Region 248..264
 FT /note="Hydrophobic sequence presumed to be an integral
 FT membrane protein anchor sequence"
 XX
 PN WO8800595-A.
 XX
 PD 28-JAN-1988.
 XX
 XX 17-JUL-1987; 87WO-AU00227.
 PR
 XX 24-MAR-1987; 87AU-0001048.
 PR
 XX (SARA-) SARAMANE PTY LTD (EPPI/).
 PA
 XX Epping RJ, Ramasamy R, Smythe JA, Anders RF, Coppel RL, Geysen HM;
 PI Saul AJ;
 PI
 XX

DR WPI; 1988-036431/05.
 DR N-PSDB; AAN80340.
 XX Merozoite surface antigen of Plasmodium falciparum -
 PT used for active immunisation or for producing monoclonal
 PT antibodies for passive immunisation, purificn. or diagnosis
 PS
 XX Claim 3; Fig 6; 59pp; English.
 PS
 XX An antigen having the AA sequence in AAP80365 or an antigenic fragment
 CC thereof is claimed. Also claimed is a recombinant DNA molecule
 CC comprising all or a portion of a nucleotide sequence which is capable of
 CC being expressed as polypeptide having the antigenicity of the above
 CC antigen or an antigen fragment. Monoclonal antibodies produced using the
 CC antigens prevent release of merozoites into the blood stream of an
 CC infected individual. The antigens can also be used for actively
 CC immunising a host against P.falciparum. The antibodies can also be used
 CC for purificn. or detection of P.falciparum.
 XX Sequence 264 AA;
 SQ

Query Match 100.0%; Score 41; DB 9; Length 264;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNTFINNA 8
 Db 27 sntfinna 34

RESULT 3
 AAR05876
 ID AAR05876 standard; protein; 272 AA.
 XX
 AC AAR05876;
 XX
 DT 11-DEC-1991 (first entry)
 XX
 XX Merozoite surface antigen-2 of the P.falciparum 3D7 isolate.
 DE
 DE MSA2; Indochina 1 isolate; malaria.
 KW
 XX Plasmodium falciparum.
 OS
 XX Key Location/Qualifiers
 FH Active-site 27..34
 FT /label= Antigenic Sequence
 FT /note="Claim 12"
 FT Active-site 207..214
 FT /label= Antigenic sequence
 FT /note="Claim 12"
 FT Active-site 221..228
 FT /label= Antigenic sequence
 FT /note="Claim 12"
 FT Region 101..108
 FT /label= Thr repeat sequence
 XX
 XX WO9002752-A.
 PN
 XX 22-MAR-1990.
 XX
 XX 12-SEP-1989; 89WO-AU00388.
 XX
 XX 24-AUG-1989; 89AU-0005962.
 PR
 PR 12-SEP-1988; 88AU-0000382.
 PR
 XX (SARA-) SARAMANE PTY LTD.
 PA
 XX Smythe JA, Anders RF, Coppel RL, Saul AJ, Jones GL;
 XX Irving DO, Dyer SL;
 XX WPI; 1990-115951/15.
 DR

DR N-PSDB; AAQ03876.
 XX Allelic variants of P falciparum merozoite surface antigen - used
 PT in prodn. of antibodies and in vaccines against malaria.
 XX
 XX Disclosure; Fig 1; 53pp; English.
 XX
 CC The merozoite surface antigen-2 may be used to generate Abs for
 CC detection and treatment of malaria, and in vaccination against the
 CC Plasmodium parasite. The isolate has conserved N- and C-terminal
 CC regions, but differs in central regions, sufficient that it may be
 CC recognised by specific probes.
 CC See also AAR05876-9.
 XX
 SQ Sequence 272 AA;

Query Match 100.0%; Score 41; DB 11; Length 272;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNTFINNA 8
 Db 27 sntfinna 34
 |||||

RESULT 4
 AAB18202
 ID AAB18202 standard; Protein; 272 AA.
 XX
 AC AAB18202;
 XX
 DT 07-NOV-2000 (first entry)
 XX
 DE Plasmodium falciparum chromosome 2 related protein SEQ ID NO:59.
 XX
 KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
 KW antimalarial; malaria; protozoacide; infection; insecticide.
 XX
 OS Plasmodium falciparum.
 XX
 PN WO200025728-A2.
 XX
 PD 11-MAY-2000.
 XX
 PF 05-NOV-1999; 99WO-US26796.
 XX
 PR 05-NOV-1998; 98US-0107131.
 XX
 PA (HOFF/) HOFFMAN S.
 PA (CARU/) CARUCCI D.
 PA (GARD/) GARDNER M.
 PA (VENT/) VENTER J C.
 XX
 PI Hoffman S, Carucci D, Gardner M, Venter JC;
 XX
 DR WPI; 2000-365347/31.
 XX
 PT Proteins encoded by chromosome 2 of the human malarial parasite,
 PT Plasmodium falciparum, useful as antimalarial vaccines and in the
 PT diagnosis of P.falciparum infection -
 XX
 PS Disclosure; Page 139; 577pp; English.
 XX
 CC The present invention describes proteins and their fragments (I) encoded
 CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
 CC Also described are: (i) nucleotide sequences (II) encoding (I); and (2)
 CC vaccines against P. falciparum infection comprising (I) or (II).
 CC (i) and (ii) are useful for the development of vaccines against
 CC P. falciparum infection. (I) and polyclonal antisera or a monoclonal
 CC antibody raised to immunogens comprising the sequences of (I), are
 CC useful in the detection of infection with P. falciparum. Furthermore,
 CC (I) (especially when they are rifins or secreted or membrane proteins)

CC can aid the identification of drugs to treat or prevent P. falciparum
 CC infection, or they can be used to identify drug resistance in
 CC P. falciparum. Sequencing of the Plasmodium chromosome 2 and the
 CC subsequent identification of proteins encoded by it will help to expand
 CC our understanding of parasite biology, a process hampered by the
 CC complexity of the parasitic lifecycle, and provide new targets for
 CC vaccine and drug development. Parasite resistance to drugs and mosquito
 CC resistance to insecticides have led to a resurgence of malaria in many
 CC parts of the world, and there is a pressing need for vaccines and new
 CC drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352 represent nucleotide
 CC and protein sequences given in the present invention, but which are not
 CC specifically mentioned within the specification.
 XX
 SQ Sequence 272 AA;

Query Match 100.0%; Score 41; DB 21; Length 272;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNTFINNA 8
 Db 27 sntfinna 34
 |||||

RESULT 5
 AAR05879
 ID AAR05879 standard; Protein; 280 AA.
 XX
 AC AAR05879;
 XX
 DT 11-DEC-1991 (first entry)
 XX
 DE Merozoite surface antigen-2 of the P.falciparum K1 isolate.
 XX
 KW MSA2; Indochina 1 isolate; malaria.
 XX
 OS Plasmodium falciparum.
 XX
 FH Key Location/Qualifiers
 FT Region 109..108
 FT /label= Repeat unit
 FT /note= "Repeated four times"
 XX
 PN WO9002752-A.
 XX
 PD 22-MAR-1990.
 XX
 PF 12-SEP-1989; 89WO-AU00388.
 XX
 PR 24-AUG-1989; 89AU-0005962.
 PR 12-SEP-1988; 88AU-0000382.
 XX
 PA (SARA-) SARAWANE PTY LTD.
 XX
 PI Smythe JA, Anders RF, Coppel RL, Saul AJ, Jones GL;
 PI Irving DO, Dyer SL;
 XX
 DR WPI; 1990-115951/15.
 DR N-PSDB; AAQ03879.
 XX
 PT Allelic variants of P falciparum merozoite surface antigen - used
 PT in prodn. of antibodies and in vaccines against malaria.
 XX
 PS Disclosure; Fig 4; 53pp; English.
 XX
 CC The merozoite surface antigen-2 may be used to generate Abs for
 CC detection and treatment of malaria, and in vaccination against the
 CC Plasmodium parasite. The isolate has conserved N- and C-terminal
 CC regions, but differs in central regions, sufficient that it may be
 CC recognised by specific probes.
 CC See also AAR05876-9.
 XX

```

SQ Sequence 280 AA;

Query Match 100.0%; Score 41; DB 11; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
Db 27 sntfinna 34

RESULT 6
AAR05877
ID AAR05877 standard; Protein; 287 AA.
AC
XX
XX
DT 11-DEC-1991 (first entry)
DE Merozite surface antigen-2 of the P.falciparum Indochina 1 isolate.
KW MSA2; 3D7 isolate; malaria.
XX
OS Plasmodium falciparum.
XX
XX Key Location/Qualifiers
FH 127..134
FT /label= Thr repeat region
XX
XX WO9002752-A.
XX
XX 22-MAR-1990.
XX
XX 12-SEP-1989; 89WO-AU00388.
XX
XX 24-AUG-1989; 89AU-0005962.
XX
XX 12-SEP-1988; 88AU-0000382.
XX
XX (SARA-) SARAMANE PTY LTD.
XX
XX Smythe JA, Anders RF, Coppel RL, Saul AJ, Jones GL;
PI Irving DO, Dyer SL;
XX
XX WPI: 1990-115951/15.
XX
XX N-PSDB; AAQ03877.
XX
XX Allelic variants of P falciparum merozoite surface antigen - used
PT in prodn. of antibodies and in vaccines against malaria.
XX
XX Disclosure; Fig 2; 53pp; English.
XX
XX The merozoite surface antigen-2 may be used to generate Abs for
CC detection and treatment of malaria, and in vaccination against the
CC Plasmodium parasite. The isolate has conserved N- and C-terminal
CC regions, but differs in central regions, sufficient that it may be
CC recognised by specific probes.
XX
XX See also AAR05876-9.
XX
XX Sequence 287 AA;

Query Match 100.0%; Score 41; DB 11; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
Db 27 sntfinna 34

RESULT 7
AAR05878
ID AAR05878 standard; Protein; 300 AA.
AC
XX
XX
DT 11-DEC-1991 (first entry)
DE Merozite surface antigen-2 of the P.falciparum MAD71 isolate.
KW MSA2; Indochina 1 isolate; malaria.
XX
OS Plasmodium falciparum.
XX
XX Key Location/Qualifiers
FH 129..136
FT /label= Thr repeat region
XX
XX WO9002752-A.
XX
XX 22-MAR-1990.
XX
XX 12-SEP-1989; 89WO-AU00388.
XX
XX 24-AUG-1989; 89AU-0005962.
XX
XX 12-SEP-1988; 88AU-0000382.
XX
XX (SARA-) SARAMANE PTY LTD.
XX
XX Smythe JA, Anders RF, Coppel RL, Saul AJ, Jones GL;
PI Irving DO, Dyer SL;
XX
XX WPI: 1990-115951/15.
XX
XX N-PSDB; AAQ03876.
XX
XX Allelic variants of P falciparum merozoite surface antigen - used
PT in prodn. of antibodies and in vaccines against malaria.
XX
XX Disclosure; Fig 1; 53pp; English.
XX
XX The merozoite surface antigen-2 may be used to generate Abs for
CC detection and treatment of malaria, and in vaccination against the
CC Plasmodium parasite. The isolate has conserved N- and C-terminal
CC regions, but differs in central regions, sufficient that it may be
CC recognised by specific probes.
XX
XX See also AAR05876-9.
XX
XX Sequence 300 AA;

Query Match 100.0%; Score 41; DB 11; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
Db 27 sntfinna 34

RESULT 8
AAY70278
ID AAY70278 standard; Protein; 350 AA.
AC
XX
XX
DT 06-JUN-2000 (first entry)
DE Recombinant vaccine CDC/NIIMALVAC-1.
XX
XX Recombinant protein; CDC/NIIMALVAC-1; multivalent; malaria; vaccine;
KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1;
KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;

```

Pfg27; antiparasitic; prevention; anti-CDC/NIIMALVAC-1 antibody;
honey bee.

Chimeric - Apis sp.
Chimeric - Clostridium tetani.
Chimeric - Plasmodium falciparum.
Key Location/Qualifiers
Peptide 1..22
/label= Melittin signal peptide
/note= "Derived from Honey bee"
Protein 23..350
/label= Mature_CDC/NIIMALVAC-1
/note= "Recombinant multivalent malarial vaccine"

WO200011179-A1.

02-MAR-2000.

19-AUG-1999; 99WO-US18869.

21-AUG-1998; 98US-0097703.

(NAIM-) NAT INST IMMUNOLOGY.
(USSH) US DEPT HEALTH & HUMAN SERVICES.

Lal AA, Shi YP, Hasnain SE;

WPI; 2000-237654/20.

N-PSDB; AA251336.

Novel recombinant protein as vaccine for treating malarial infection
comprises antigenic peptides obtained from different stages of
plasmodium falciparum life cycle -

Claim 3; Page 43-44; 52pp; English.

The present sequence is that of recombinant protein CDC/NIIMALVAC-1,
which is a multivalent, multistage malarial vaccine. The recombinant
protein comprises, melittin signal peptide, (His)6 tag, T-cell epitope
from tetanus toxoid and 21 antigenic epitopes from circumsporozoite
protein (CSP), sporozoite surface protein-2 (SSP-2), liver stage
antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1), MSP-2, apical
membrane antigen-1 (AMA-1), erythrocyte binding antigen-175 (EBA-175),
roptery associated protein-1 (RAP-1) and gamete specific antigen, Pfg27.
These epitopes were obtained at different stages of the life cycle of
Plasmodium falciparum. CDC/NIIMALVAC-1 vaccine has antiparasitic
activity and can be used for treatment and prevention of malarial
infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for detecting
P. falciparum in biological samples.

Sequence 350 AA;

Query Match 100.0%; Score 41; DB 21; Length 350;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8

Db 207 sntfinna 214

RESULT 9

AA03985

ID AAR03985 standard; Peptide; 17 AA.

AC AAR03985;

DT 11-DEC-1991 (first entry)

DE Antigenic peptide fragment of MSA2 of P.falciparum Indochina 1 isolate.

XX

KW MSA2; 3D7 isolate; malaria.
XX Plasmodium falciparum.
OS WO9002752-A.
PN 22-MAR-1990.
XX 12-SEP-1989; 89WO-AU00388.
XX 24-AUG-1989; 89AU-0005962.
PR 12-SEP-1988; 88AU-0000382.
XX (SARA-) SARAMANE PTY LTD.
XX Smythe JA, Anders RF, Coppel RL, Saul AJ, Jones GL;
PI Irving DO, Dyer SL;
PI WPI; 1990-115951/15.
XX Allelic variants of P falciparum merozoite surface antigen - used
in prodn. of antibodies and in vaccines against malaria.
XX Disclosure; Page 12; 53pp; English.
XX The merozoite surface antigen-2 may be used to generate Abs for
detection and treatment of malaria, and in vaccination against the
Plasmodium parasite. The isolate has conserved N- and C-terminal
regions, but differs in central regions, sufficient that it may be
recognised by specific probes.
XX See also AAR05876-9.
XX Sequence 17 AA;

Query Match 90.2%; Score 37; DB 11; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINN 7

Db 11 sntfinn 17

RESULT 10

AA04944

ID AAR04944 standard; protein; 469 AA.

XX AAR04944;

DT 02-OCT-1990 (first entry)

DE Equine neuraminidase N7 (EIV-A1).

KW Recombinant vaccines; equine influenza virus; haemagglutinin; H7;
neuraminidase; N7.

XX Equine influenza virus.

XX US920213-A.

PD 24-APR-1990.

XX 21-JUL-1986; 86US-0888250.

PR 21-JUL-1986; 86US-0888250, US-747020.

XX (BIOT-) BIOTECH RES PARTNER.

XX Dale B, Cordell B;

XX WPI; 1990-163647/21.

DR N-PSDB; AAQ04598.

XX Recombinant vaccines against equine influenza virus - produced using DNA
PT sequences encoding haemagglutinin and neuraminidase glycoprotein(s).
XX
PS Disclosure; : 27pp; English.
XX
CC H7 and N7 are virally encoded glycoproteins which confer the
CC immunological characteristics of influenza virions. There are nine
CC neuraminidase (NA) subtypes and twelve haemagglutinin (HA) subtypes.
CC The strain carrying H7N7 glycoproteins is designated equine influenza
CC virus (EVI)-A1. The cDNA sequences encoding these will be useful in the
CC construction of diagnostic probes for the disease and of probes for
CC obtaining new cDNAs of the mutated form of the virus. Recombinant
CC vaccines are produced.
CC See also AAQ04596-Q04599.
XX
XX Sequence 469 AA;
SQ

Query Match 82.9%; Score 34; DB 11; Length 469;
Best Local Similarity 85.7%; Pred. No. 67;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NTFINNA 8
Db 61 ntyinna 67

RESULT 11
AAP70712
ID AAP70712 standard; protein; 470 AA.
AC
XX
XX AAP70712;
DT 08-MAR-1991 (first entry)
XX
DE Equine influenza virus strain H7N7 (EIV-A1) neuraminidase N7 gene
DE product.
XX
XX HA; vaccine; Vaccina.
XX
XX Equine influenza virus.
XX
XX WO8607593-A.
XX
XX 31-DEC-1986.
XX
XX 20-JUN-1986; 86WO-US01343.
XX
XX 20-JUN-1985; 85US-0747020.
XX
XX (BIOT-) BIOTECHN RES PARTNE.
XX
XX Dale B, Cordell B;
XX
XX WPI; 1987-007191/01.
XX
XX N-PSDB; AAN71068.
XX
XX preventing equine influenza virus infection - using recombinant
XX vaccines produced using DNA sequences encoding haemagglutinin and
XX neuraminidase glyco:proteins
XX
XX Disclosure; Fig 3; 63pp; English.
XX
XX Peptides derived from the haemagglutinin H7 and H3 and neuraminidase
XX N7 and N8 genes may be used to derive antigenic peptides useful in
XX vaccination against equine influenza virus infection.
XX Abs raised to the peptides may be used in diagnosis of the infection
XX and construction of probes to mutated forms of the virus.
XX
XX Sequence 470 AA;

Query Match 82.9%; Score 34; DB 8; Length 470;
Best Local Similarity 85.7%; Pred. No. 67;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NTFINNA 8
Db 61 ntyinna 67

RESULT 12
AAB18218
ID AAB18218 standard; Protein; 2010 AA.
XX
XX AAB18218;
XX
XX 07-NOV-2000 (first entry)
XX
XX Plasmodium falciparum chromosome 2 related protein SEQ ID NO:75.
XX
XX Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
XX antimalarial; malaria; protozoacide; infection; insecticide.
XX
XX Plasmodium falciparum.
XX
XX WO200025728-A2.
XX
XX 11-MAY-2000.
XX
XX 05-NOV-1999; 99WO-US26796.
XX
XX 05-NOV-1998; 98US-0107131.
XX
XX (HOFF/) HOFFMAN S.
XX (CARU/) CARUCCI D.
XX (GARD/) GARDNER M.
XX (VENT/) VENTER J C.
XX
XX Hoffman S, Carucci D, Gardner M, Venter JC;
XX
XX WPI; 2000-365347/31;
XX
XX Proteins encoded by chromosome 2 of the human malarial parasite,
XX Plasmodium falciparum, useful as antimalarial vaccines and in the
XX diagnosis of P.falciparum infection -
XX
XX Disclosure; Page 177-182; 577pp; English.
XX
XX The present invention describes proteins and their fragments (I) encoded
XX by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
XX Also described are: (I) nucleotide sequences (II) encoding (I); and (2)
XX vaccines against P. falciparum infection comprising (I) or (II).
XX (I) and (II) are useful for the development of vaccines against
XX P. falciparum infection. (I) and polyclonal antisera or a monoclonal
XX antibody raised to immunogens comprising the sequences of (I), are
XX useful in the detection of infection with P. falciparum. Furthermore,
XX (I) (especially when they are rifins or secreted or membrane proteins)
XX can aid the identification of drugs to treat or prevent P. falciparum
XX infection, or they can be used to identify drug resistance in
XX P. falciparum. Sequencing of the Plasmodium chromosome 2 and the
XX subsequent identification of proteins encoded by it will help to expand
XX our understanding of parasite biology, a process hampered by the
XX complexity of the parasitic lifecycle, and provide new targets for
XX vaccine and drug development. Parasite resistance to drugs and mosquito
XX resistance to insecticides have led to a resurgence of malaria in many
XX parts of the world, and there is a pressing need for vaccines and new
XX drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352 represent nucleotide
XX and protein sequences given in the present invention, but which are not
XX specifically mentioned within the specification.
XX
XX Sequence 2010 AA;

Query Match 80.5%; Score 33; DB 21; Length 2010;

Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NTFINN 7
Db 735 ntfinn 740
|||||

RESULT 13

AAR46082
ID AAR46082 standard; Protein; 68 AA.

XX
AC AAR46082;

XX
DT 19-OCT-1994 (first entry)

XX
DE DEAD ATP helicase like protein.

XX
KW Human cDNA; library; enzyme; protein.

XX
OS Homo sapiens.

XX
PN WO9403599-A.

XX
PD 17-FEB-1994.

XX
PF 04-AUG-1993; 93WO-JP01095.

XX
PR 04-AUG-1992; 92JP-0208077.

XX
PR 13-NOV-1992; 92JP-0327619.

XX
PR 26-FEB-1993; 93JP-0061431.

XX
PA (SAGA) SAGAMI CHEM RES CENTRE.

XX
PI Iwahori A, Kato S, Kato T, Kim N, Oh S, Sekine S;

XX
DR WPI; 1994-065688/08.

XX
DR N-PSDB; AAQ57421.

PT cDNA of human origin and proteins coded by it - which may be

PT expressed by in vivo or in vitro translation using sense RNA or

PT antisense DNA corresponding to the cDNA.

XX
PS Claim 1; Page 37; 167pp; Japanese.

XX
CC mRNA expressed in human fibrosarcoma cell line HT-1080 was

XX
CC isolated and used to construct a cDNA library using vector

XX
CC pK1. Clone HP00038 encoding DEAD ATP helicase-like protein

XX
SQ Sequence 68 AA;

Query Match

Best Local Similarity 78.0%; Score 32; DB 15; Length 68;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTFINNA 8
Db 37 attfinna 44
: |||||

RESULT 14

AAG98434
ID AAG98434 standard; Protein; 216 AA.

XX
AC AAG98434;

XX
DT 25-SEP-2001 (first entry)

XX
DE Papio hamadryas olfactory receptor 3.

XX
KW Olfactory receptor; primate; mouse; human; food processing industry;

KW aromas; perfumery; toxic substance.

XX
OS Papio hamadryas.

XX
PN WO200146262-A2.

XX
PD 28-JUN-2001.

XX
PF 22-DEC-2000; 2000WO-IB02017.

XX
PR 22-DEC-1999; 99US-0171746.

XX
PR 21-DEC-2000; 2000US-0747155.

XX
PA (CNRS) CNRS CENT NAT RECH SCI.

XX
PI Rouquier S, Giorgi D;

XX
DR WPI; 2001-381911/40.

XX
DR N-PSDB; AAH84126.

XX
PT Nucleic acids encoding primate and murine olfactory receptors, useful

XX
PS Claim 3; Page 42-43; 482pp; English.

XX
CC The invention relates to olfactory receptors (AAG98432-AAG98609) and the

XX
CC genes encoding them (AAH83879-AAH84131) including pseudogenes of 10

XX
CC primate species, mouse and human. The nucleic acids and receptors may be

XX
CC used in the food processing industry (e.g. for the detection of aromas,

XX
CC quality control and sample analysis), in perfumery (e.g. for the analysis

XX
CC or comparison of perfumes) and in the environment (e.g. for the detection

XX
SQ Sequence 216 AA;

Query Match

Best Local Similarity 78.0%; Score 32; DB 22; Length 216;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTFINN 7
Db 123 sdtfinn 129
: |||||

RESULT 15

AAG98448
ID AAG98448 standard; Protein; 216 AA.

XX
AC AAG98448;

XX
DT 25-SEP-2001 (first entry)

XX
DE Gorilla olfactory receptor 5.

XX
KW Olfactory receptor; primate; mouse; human; food processing industry;

XX
KW aromas; perfumery; toxic substance.

XX
OS Gorilla gorilla.

XX
PN WO200146262-A2.

XX
PD 28-JUN-2001.

XX
PF 22-DEC-2000; 2000WO-IB02017.

XX
PR 22-DEC-1999; 99US-0171746.

XX
PR 21-DEC-2000; 2000US-0747155.

XX
PA (CNRS) CNRS CENT NAT RECH SCI.

XX
PI Rouquier S, Giorgi D;

DR WPI; 2001-381911/40.
DR N-PSDB; AAH84131.
XX
XX Nucleic acids encoding primate and murine olfactory receptors, useful
PT for analysis odours e.g. in food processing and perfumery -
XX
XX Claim 3; Page 89-90; 482pp; English.
PS
XX
XX The invention relates to olfactory receptors (AAC98432-AAG98609) and the
CC genes encoding them (AAH83879-AAH84131) including pseudogenes of 10
CC primate species, mouse and human. The nucleic acids and receptors may be
CC used in the food processing industry (e.g. for the detection of aromas,
CC quality control and sample analysis), in perfumery (e.g. for the analysis
CC or comparison of perfumes) and in the environment (e.g. for the detection
CC of toxic substances and/or trapping of odours).
XX
SQ Sequence 216 AA;

Query Match 78.0%; Score 32; DB 22; Length 216;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINN 7
|:|||||
Db 123 sdtfinn 129

Search completed: January 29, 2002, 10:21:47
Job time: 423 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:24:06 : Search time 133.18 seconds
(without alignments)
1.352 Million cell updates/sec

Title: US-09-763-397A-14
Perfect score: 41
Sequence: 1 SNTFINNA 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	75.6	499	2	US-09-032-315-3
2	31	75.6	499	2	US-08-993-318A-3
3	31	75.6	499	4	US-09-399-886-3
4	31	75.6	499	4	US-09-396-260-3
5	31	75.6	499	4	US-09-576-281-3
6	31	75.6	519	1	US-08-462-484-4
7	31	75.6	519	1	US-08-441-147-4
8	31	75.6	519	5	PCT-US95-07536-4
9	30	73.2	51	1	US-08-445-909A-18
10	30	73.2	309	3	US-08-988-876-5
11	30	73.2	527	1	US-08-462-484-10
12	30	73.2	527	1	US-08-441-147-10
13	30	73.2	527	5	PCT-US95-07536-10
14	29	70.7	905	1	US-08-072-574-2
15	29	70.7	906	1	US-08-486-270-2
16	29	70.7	906	3	US-08-367-264-2
17	29	70.7	906	5	PCT-US91-09422-17
18	29	70.7	1194	4	US-08-538-526-1
19	29	70.7	1199	1	US-08-041-538-2
20	29	70.7	1199	1	US-08-463-642-2
21	29	70.7	1199	1	US-08-455-602-2
22	29	70.7	1199	2	US-08-465-157-2
23	29	70.7	1199	5	PCT-US91-09422-2
24	29	70.7	1219	2	US-08-687-289A-6
25	28	68.3	123	4	US-09-124-900-10
26	28	68.3	206	3	US-08-388-353-644
27	28	68.3	206	3	US-08-488-551B-644

28 68.3 260 4 US-08-081-929-10 Sequence 10, Appl
29 68.3 270 1 US-08-452-052-1 Sequence 1, Appl
30 68.3 373 1 US-08-015-986A-14 Sequence 14, Appl
31 68.3 373 2 US-08-446-363-14 Sequence 14, Appl
32 68.3 2308 1 US-08-015-973-1 Sequence 1, Appl
33 68.3 2308 4 US-08-081-929-2 Sequence 1, Appl
34 68.3 2308 4 US-08-448-164-1 Sequence 2, Appl
35 68.3 136 4 US-09-456-830-9 Sequence 9, Appl
36 68.3 136 4 US-09-456-830-17 Sequence 17, Appl
37 65.9 285 4 US-08-992-035A-1 Sequence 1, Appl
38 65.9 355 4 US-09-308-595C-10 Sequence 10, Appl
39 65.9 392 1 US-08-423-441-2 Sequence 2, Appl
40 65.9 459 1 US-09-080-983-5 Sequence 5, Appl
41 65.9 519 1 US-08-462-949-2 Sequence 2, Appl
42 65.9 519 1 US-08-023-764B-2 Sequence 2, Appl
43 65.9 579 1 US-08-448-166A-8 Sequence 8, Appl
44 65.9 612 3 US-09-293-186-16 Sequence 16, Appl
45 65.9 645 2 US-08-592-126-144 Sequence 144, Appl

ALIGNMENTS

RESULT 1

US-09-032-315-3
: Sequence 3, Application US/09032315
: Patent No. 5985818
: GENERAL INFORMATION:
: APPLICANT: Svendsen, Allan
: APPLICANT: Xu, Feng
: TITLE OF INVENTION: LACCASE MUTANTS
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 5985818o No. 5985818disk of No. 5985818th America, Inc.
: STREET: 405 Lexington Avenue
: CITY: New York
: STATE: NY
: COUNTRY: USA
: ZIP: 10174
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/032,315
: FILING DATE: 27-FEB-1998
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Rozek, Carol
: REGISTRATION NUMBER: 36,993
: REFERENCE/DOCKET NUMBER: 5200.200-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-867-0123
: TELEFAX: 212-878-9655
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 499 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-09-032-315-3

Query Match 75.6%; Score 31; DB 2; Length 499;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNTFINNA 8

Db 335 TNFFINNA 342

RESULT 2
US-08-993-318A-3
; Sequence 3, Application US/08993318A
; Patent No. 5998353
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Anders
; APPLICANT: Svendsen, Allan
; APPLICANT: Schneider, Palle
; APPLICANT: Rasmussen, Grethe
; APPLICANT: Cherry, Joel
; TITLE OF INVENTION: LACCASE MUTANTS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5998353o No. 5998353disk of No. 5998353th America
; STREET: 405 Lexington Avenue
; CITY: New York
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,318A
; FILING DATE: December 18, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gregg, Valeta A.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 5032.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 499 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-993-318A-3

Query Match 75.6%; Score 31; DB 2; Length 499;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNTFFINNA 8
; :| |||||
DB 335 TNEFFINNA 342

RESULT 3
US-09-399-886-3
; Sequence 3, Application US/09399886
; Patent No. 6140092
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Anders
; APPLICANT: Svendsen, Allan
; APPLICANT: Schneider, Palle
; APPLICANT: Rasmussen, Grethe
; APPLICANT: Cherry, Joel
; TITLE OF INVENTION: LACCASE MUTANTS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6140092o No. 6140092disk of No. 6140092th America
; STREET: 405 Lexington Avenue
; CITY: New York
; COUNTRY: USA
; ZIP: 10174

Query Match 75.6%; Score 31; DB 2; Length 499;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNTFFINNA 8
; :| |||||
DB 335 TNEFFINNA 342

RESULT 4
US-09-396-260-3
; Sequence 3, Application US/09396260
; Patent No. 6184015
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Xu, Feng
; TITLE OF INVENTION: LACCASE MUTANTS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6184015o No. 6184015disk of No. 6184015th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/396,260
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,315
; FILING DATE: 27-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Rozek, Carol
; REGISTRATION NUMBER: 36,993
; REFERENCE/DOCKET NUMBER: 5200.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 499 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-396-260-3

Query Match 75.6%; Score 31; DB 4; Length 499;
Best Local Similarity 75.0%; Pred. NO. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SNTFINNA 8
: | | | | | | |
Db 335 TNFFINNA 342

RESULT 5
US-09-576-281-3
Sequence 3, Application US/09576281
Patent No. 6277611
GENERAL INFORMATION:

APPLICANT: Pedersen, Anders
APPLICANT: Svendsen, Allan
APPLICANT: Schneider, Palle
APPLICANT: Rasmussen, Grethe
APPLICANT: Cherry, Joel
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 6277611 No. 6277611disk of No. 6277611th America
STREET: 405 Lexington Avenue
CITY: New York
COUNTRY: USA
ZIP: 10174

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/576,281
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,318
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeta A.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 5032,200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 499 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-576-281-3

Query Match 75.6%; Score 31; DB 4; Length 499;
Best Local Similarity 75.0%; Pred. NO. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SNTFINNA 8

Db 335 TNFFINNA 342
: | | | | | | |

RESULT 6
US-08-462-484-4
Sequence 4, Application US/08462484
Patent No. 5667531
GENERAL INFORMATION:

APPLICANT: Yaver, Debbie Sue
APPLICANT: Xu, Feng
APPLICANT: Dalboge, Henrik
APPLICANT: Schneider, Palle
APPLICANT: Aaslyng, Dorrit A.
TITLE OF INVENTION: PURIFIED POLYPORUS LACCASES AND
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING SAME
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 5667531 No. 5667531disk of No. 5667531th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,484
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/441,147
FILING DATE: 15-MAY-1995
ATTORNEY/AGENT INFORMATION:

NAME: Lowney, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4185,010-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 878 9655
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 519 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-484-4

Query Match 75.6%; Score 31; DB 1; Length 519;
Best Local Similarity 75.0%; Pred. NO. 1.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SNTFINNA 8
: | | | | | | |
Db 355 TNFFINNA 362

RESULT 7
US-08-441-147-4
Sequence 4, Application US/08441147
Patent No. 5770418
GENERAL INFORMATION:

APPLICANT: Yaver, Debbie Sue
APPLICANT: Xu, Feng
APPLICANT: Dalboge, Henrik
APPLICANT: Schneider, Palle
APPLICANT: Aaslyng, Dorrit A.
TITLE OF INVENTION: PURIFIED POLYPORUS LACCASES AND
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING SAME

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 57704180 No. 5770418disk of No. 5770418th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,147
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lowney, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4185.010-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 878 9655
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 519 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-441-147-4

Query Match 75.6%; Score 31; DB 1; Length 519;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
Db 355 TNFFINNA 362

RESULT 8
PCT-US95-07536-4
Sequence 4, Application PC/TUS9507536
GENERAL INFORMATION:
APPLICANT:
APPLICANT:
TITLE OF INVENTION: PURIFIED POLYPORUS LACCASES AND
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING SAME
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Novo Nordisk of North America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07536
FILING DATE: 15-June-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/265,534
FILING DATE: 24-June-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowney, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4185.204-WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123

TELEFAX: 212 878 9655
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 519 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-07536-4

Query Match 75.6%; Score 31; DB 5; Length 519;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
Db 355 TNFFINNA 362

RESULT 9
US-08-445-909A-18
Sequence 18, Application US/08445909A
Patent No. 5772996
GENERAL INFORMATION:
APPLICANT: ATKINSON, Antony
TITLE OF INVENTION: Pharmaceutical Compositions
TITLE OF INVENTION: Containing Superoxide Dismutase from Bacillus
TITLE OF INVENTION: Stearothermophilus and Bacillus Caldotenax
NUMBER OF SEQUENCES: 29
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445,909A
FILING DATE: 22-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/978,697
FILING DATE: 02-FEB-1993
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 51 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-445-909A-18

Query Match 73.2%; Score 30; DB 1; Length 51;
Best Local Similarity 83.3%; Pred. No. 23;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFIN 6
Db 34 SNTFVN 39

RESULT 10
US-08-988-876-5
Sequence 5, Application US/08988876
Patent No. 6063596
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
TITLE OF INVENTION: G PROTEIN COUPLED RECEPTORS ASSOCIATED
TITLE OF INVENTION: WITH IMMUNE RESPONSE
NUMBER OF SEQUENCES: 9

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Incyte Pharmaceuticals, Inc.
;; STREET: 3174 Porter Drive
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/988,876
;; FILING DATE: Herewith
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Billings, Lucy J.
;; REGISTRATION NUMBER: 36,749
;; REFERENCE/DOCKET NUMBER: PF-0441 US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650-855-0355
;; TELEFAX: 650-845-4166
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 309 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: GenBank
;; CLONE: 1314667
;; US-08-988-876-5

Query Match 73.2%; Score 30; DB 3; Length 309;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINN 7
|:|:|:|
Db 190 SDFLNN 196

RESULT 11
US-08-462-484-10
; Sequence 10, Application US/08462484
; Patent No. 5667531
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Xu, Feng
; APPLICANT: Dalboge, Henrik
; APPLICANT: Schneider, Palle
; APPLICANT: Asalvng, Dorrit A.
; TITLE OF INVENTION: PURIFIED POLYPORUS LACCASES AND
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5667531o No. 5667531disk of No. 5667531th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/462,484
;; FILING DATE: 05-JUN-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/441,147
;; FILING DATE: 15-MAY-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lowney, Karen A.
;; REGISTRATION NUMBER: 31,274
;; REFERENCE/DOCKET NUMBER: 4185.010-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212 867 0123
;; TELEFAX: 212 878 9655
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 527 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; ORIGINAL SOURCE:
;; ORGANISM: Polyporus pinsitus
;; US-08-462-484-10

Query Match 73.2%; Score 30; DB 1; Length 527;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNTFINN 7
|:|:|:|
Db 360 SNEFINN 366

RESULT 12
US-08-441-147-10
; Sequence 10, Application US/08441147
; Patent No. 5770418
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Xu, Feng
; APPLICANT: Dalboge, Henrik
; APPLICANT: Schneider, Palle
; APPLICANT: Asalvng, Dorrit A.
; TITLE OF INVENTION: PURIFIED POLYPORUS LACCASES AND
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5770418o No. 5770418disk of No. 5770418th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/441,147
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4185.010-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 878 9655
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:

LENGTH: 527 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Polyporus pinsitus
US-08-441-147-10

Query Match 73.2%; Score 30; DB 1; Length 527;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNTFINN 7
II IIII
DB 360 SNTFINN 366

RESULT 13
PCT-US95-07536-10
Sequence 10, Application PC/TUS9507536
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: PURIFIED POLYPORUS LACCASES AND
NUCLEIC ACIDS ENCODING SAME
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Novo Nordisk of North America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
COUNTRY: U.S.A.
ZIP: 10174-6401

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07536
FILING DATE: 15-June-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/265,534
FILING DATE: 24-June-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowney, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4185,204-WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 878 9655

INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Polyporus pinsitus
PCT-US95-07536-10

Query Match 73.2%; Score 30; DB 5; Length 527;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNTFINN 7
II IIII
DB 360 SNTFINN 366

RESULT 14

US-08-072-574-2
Sequence 2, Application US/08072574
Patent No. 5521297
GENERAL INFORMATION:
APPLICANT: Daggett, Lorrie
APPLICANT: Ellis, Steven B.
APPLICANT: Liaw, Chen
APPLICANT: Pontsler, Aaron
TITLE OF INVENTION: HUMAN METABOTROPIC GLUTAMATE RECEPTORS,
NUCLEIC ACIDS ENCODING SAME AND USES THEREOF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 00719
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/072,574
FILING DATE: 19930604
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9383
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213-622-7700
TELEFAX: 213-489-4210
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 905 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-072-574-2

Query Match 70.7%; Score 29; DB 1; Length 905;
Best Local Similarity 83.3%; Pred. No. 6.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFIN 6
IIIIII
DB 868 SNTFIN 873

RESULT 15
US-08-486-270-2
Sequence 2, Application US/08486270
Patent No. 5807689
GENERAL INFORMATION:
APPLICANT: Daggett, Lorrie
APPLICANT: Ellis, Steven B.
APPLICANT: Liaw, Chen
APPLICANT: Pontsler, Aaron
APPLICANT: Johnson, Edwin C.
APPLICANT: Hess, Stephen D.
TITLE OF INVENTION: HUMAN METABOTROPIC GLUTAMATE RECEPTORS,
NUCLEIC ACIDS ENCODING SAME AND USES THEREOF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,270
FILING DATE: 02-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/072,574
FILING DATE: 04-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: EP41 9772
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4737
TELEFAX: 619-546-9392
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 906 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-486-270-2

Query Match 70.7%; Score 29; DB 1; Length 906;
Best Local Similarity 83.3%; Pred. No. 6.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SNTFIN 6
Db 869 SNTFLN 874

Search completed: January 29, 2002, 10:24:07
Job time: 513 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:26:42 ; Search time 144.96 seconds
(without alignments)
4.204 Million cell updates/sec

Title: US-09-763-397A-14
Perfect score: 41
Sequence: 1 SNTFINNA 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	256	2 S39311	merozoite surface
2	41	100.0	264	2 A31818	merozoite 45K surf
3	41	100.0	264	2 S53367	merozoite surface
4	41	100.0	272	2 G71618	merozoite surface
5	41	100.0	274	2 A45632	merozoite surface
6	41	100.0	278	2 S39310	merozoite surface
7	41	100.0	280	2 C39112	merozoite 45K surf
8	41	100.0	286	2 B45632	merozoite surface
9	41	100.0	287	2 B39615	merozoite 45K surf
10	41	100.0	300	2 A39112	merozoite 45K surf
11	41	100.0	302	2 A39615	merozoite 45K surf
12	41	100.0	347	2 B39112	merozoite 45K surf
13	37	90.2	202	2 S78382	ribosomal protein
14	34	82.9	469	2 NM1VEA	exo-alpha-stilidas
15	33	80.5	378	2 F56653	virion protein - h
16	33	80.5	436	2 S48399	hypothetical prote
17	33	80.5	555	2 T44010	virion protein (im
18	33	80.5	2010	2 B71616	phosphatase (acid
19	33	80.5	3036	2 T18995	hypothetical prote
20	32	78.0	157	2 S58004	probable olfactory
21	32	78.0	157	2 S57995	probable olfactory
22	32	78.0	321	2 T19259	hypothetical prote
23	32	78.0	321	2 T26153	hypothetical prote
24	32	78.0	351	1 VVVP28	coat protein VP2 -
25	32	78.0	351	1 VVVPAS	coat protein VP2 -
26	32	78.0	614	2 S45053	membrane protein E
27	32	78.0	644	2 T16136	hypothetical prote
28	32	78.0	776	2 T50085	hypothetical colle
29	32	78.0	1645	2 H85554	hypothetical prote

```

30 31 75.6 88 2 T30366 occlusion-derived
31 31 75.6 110 2 S61540 ribosomal protein
32 31 75.6 180 2 B45613 surface antigen FU
33 31 75.6 253 2 T40687 hypothetical prote
34 31 75.6 337 2 T52589 ribose-phosphate p
35 31 75.6 366 2 E96773 probable lipase/ac
36 31 75.6 380 2 T18509 hypothetical prote
37 31 75.6 507 2 F71055 hypothetical prote
38 31 75.6 511 2 H86684 prophage p11 prote
39 31 75.6 511 2 C86798 prophage p13 prote
40 31 75.6 558 2 S83148 probable membrane
41 31 75.6 567 2 S58750 NAH dehydrogenase
42 31 75.6 573 2 S50661 hypothetical prote
43 31 75.6 626 2 D70178 PTS system, fructo
44 31 75.6 700 2 S57194 calpain (EC 3.4.22
45 31 75.6 1035 2 T16588 hypothetical prote

```

ALIGNMENTS

RESULT 1

S39311
merozoite surface antigen - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 17-Nov-2000
C:Accession: S39311
R:Ramasamy, R.; Ranasinghe, C.
submitted to the EMBL data library, November 1993
A:Description: Cycle ds DNA sequencing of a malaria parasite protein from infected bl
A:Reference number: S39310
A:Accession: S39311
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-256 <RAM>
A:Cross-references: EMBL:X76298; NID:G434997; PID:G836640
C:Superfamily: Epstein-Barr virus nuclear antigen

Query Match 100.0%; Score 41; DB 2; Length 256;
Best Local Similarity 100.0%; Pred. No. 0.87;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 SNTFINNA 8
Db 27 SNTFINNA 34

```

RESULT 2

A31818
merozoite 45K surface antigen precursor - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 17-Nov-2000
C:Accession: A31818
R:Smythe, J.A.; Coppel, R.L.; Brown, G.V.; Ramasamy, R.; Kemp, D.J.; Anders, R.F.
Proc. Natl. Acad. Sci. U.S.A. 85, 5195-5199, 1988
A:Title: Identification of two integral membrane proteins of Plasmodium falciparum.
A:Reference number: A31818; MUID:88276924
A:Accession: A31818
A:Molecule type: mRNA
A:Residues: 1-264 <SMY>
A:Cross-references: CB:J03828; NID:G160352; PID:G160353
C:Superfamily: Epstein-Barr virus nuclear antigen
C:Keywords: surface antigen
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-264/Product: merozoite 45K surface antigen #status predicted <MAT>

Query Match 100.0%; Score 41; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 0.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 SNTFINNA 8

```

```

Db      27  SNTFINNA 34
|||||||
RESULT  3
Merozoite surface antigen 2 - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 15-Jul-1995 #sequence_revision-01-Sep-1995 #text_change 17-Nov-2000
C:Accession: S55367
R:Chauhan, V.S.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55367
A:Accession: S55367
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <CHA>
A:Cross-references: EMBL:X87249; NID:g854367; PID:g854368
C:Superfamily: Epstein-Barr virus nuclear antigen
C:Keywords: surface antigen

Query Match      100.0%; Score 41; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 0.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  SNTFINNA 8
    |||||
Db   27  SNTFINNA 34

RESULT  4
Merozoite surface antigen MSP-2 PFB0300c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 17-Nov-2000
C:Accession: G71618; A44950
R:Gardner, M.J.; Tetzelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, E.V.;
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743
A:Accession: G71618
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-272 <GAR>
A:Cross-references: GB:AE001385; GB:AE001362; NID:g3845143; PIDN:AAC71849.1; PID:g384514
A:Experimental source: clone 3D7
R:Smythe, J.A.; Peterson, M.G.; Coppel, R.L.; Saul, A.J.; Kemp, D.J.; Anders, R.F.
Mol. Biochem. Parasitol. 39, 227-234, 1990
A:Title: Structural diversity in the 45-kilodalton merozoite surface antigen of Plasmodium
A:Reference number: A44950; MUID:90205972
A:Accession: A44950
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-54,'T',56-272 <SMY>
A:Cross-references: GB:M28891; NID:g160458; PID:g160459
C:Genetics:
A:Gene: PFB0300c
C:Superfamily: Epstein-Barr virus nuclear antigen
C:Keywords: surface antigen

Query Match      100.0%; Score 41; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.93;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  SNTFINNA 8
    |||||
Db   27  SNTFINNA 34

RESULT  5
Merozoite surface antigen 2 - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 17-Nov-2000
C:Accession: A45632
R:Marshall, V.M.; Coppel, R.L.; Anders, R.F.; Kemp, D.J.
Mol. Biochem. Parasitol. 50, 181-184, 1992
A:Title: Two novel alleles within subfamilies of the merozoite surface antigen 2 (MSA
A:Reference number: A45632; MUID:92178286
A:Contents: K1916
A:Accession: A45632
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-274 <MAR>
A:Cross-references: GB:M73810; NID:g160484; PID:g160485
A:Note: sequence extracted from NCBI backbone (NCBIN:85252, NCBIP:85257)
C:Superfamily: Epstein-Barr virus nuclear antigen
C:Keywords: surface antigen

Query Match      100.0%; Score 41; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 0.93;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  SNTFINNA 8
    |||||
Db   27  SNTFINNA 34

RESULT  6
Merozoite surface antigen - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Nov-2000
C:Accession: S39310
R:Ramassamy, R.; Ranasinghe, C.
submitted to the EMBL Data Library, November 1993
A:Description: Cycle ds DNA sequencing of a malaria parasite protein from infected bl
A:Reference number: S39310
A:Accession: S39310
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-278 <RAM>
A:Cross-references: EMBL:X76087; NID:g434996; PID:g836639
C:Superfamily: Epstein-Barr virus nuclear antigen
C:Keywords: surface antigen

Query Match      100.0%; Score 41; DB 2; Length 278;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  SNTFINNA 8
    |||||
Db   27  SNTFINNA 34

RESULT  7
Merozoite 45K surface antigen precursor - malaria parasite (Plasmodium falciparum) (1
C:Species: Plasmodium falciparum
C>Date: 30-Aug-1991 #sequence_revision 30-Aug-1991 #text_change 17-Nov-2000
C:Accession: C39112
R:Smythe, J.A.; Coppel, R.L.; Day, K.P.; Martin, R.K.; Oduola, A.M.J.; Kemp, D.J.; An
Proc. Natl. Acad. Sci. U.S.A. 88, 1751-1755, 1991
A:Title: Structural diversity in the Plasmodium falciparum merozoite surface antigen
A:Reference number: A39112; MUID:91156685
A:Accession: C39112
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-280 <SMY>
A:Cross-references: GB:M59768
C:Superfamily: Epstein-Barr virus nuclear antigen

```

C;Keywords: surface antigen

Query Match 100.0%; Score 41; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.95; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

Oy 1 SNTFINNA 8
| | | | | | | |
Db 27 SNTFINNA 34

RESULT 8

B45632
merozoite surface antigen 2 - malaria parasite (Plasmodium falciparum)

C;Species: Plasmodium falciparum
C;Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 17-Nov-2000

C;Accession: B45632

R;Marshall, V.M.; Coppel, R.L.; Anders, R.F.; Kemp, D.J.

Mol. Biochem. Parasitol. 50, 181-184, 1992

A;Title: Two novel alleles within subfamilies of the merozoite surface antigen 2 (MSA-2)

A;Reference number: A45632; MUID:92178286

A;Accession: B45632

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-286 <MAR>

A;Experimental source: isolate 311

A;Note: sequence extracted from NCBI backbone (NCBIN:85255, NCBIP:85259)

C;Superfamily: Epstein-Barr virus nuclear antigen

C;Keywords: surface antigen

Query Match 100.0%; Score 41; DB 2; Length 286;
Best Local Similarity 100.0%; Pred. No. 0.97; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

Oy 1 SNTFINNA 8
| | | | | | | |
Db 27 SNTFINNA 34

RESULT 9

B39615
merozoite 45K surface antigen precursor - malaria parasite (Plasmodium falciparum)

N;Alternate names: membrane antigen pF7

C;Species: Plasmodium falciparum

C;Date: 13-Sep-1991 #sequence_revision 13-Sep-1991 #text_change 01-Dec-2000

C;Accession: B39615; A36018; B44950; A45613

R;Fenton, B.; Clark, J.T.; Khan, C.M.A.; Robinson, J.V.; Walliker, D.; Ridley, R.; Scaife

Mol. Cell. Biol. 11, 963-971, 1991

A;Title: Structural and antigenic polymorphism of the 35- to 48-kilodalton merozoite sur

A;Reference number: A39615; MUID:91117264

A;Accession: B39615

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-287 <FEN>

A;Cross-references: EMBL:X53833

A;Note: clone T9-94

R;Elliot, J.F.; Albrecht, G.R.; Gilladoga, A.; Handunnetti, S.M.; Neequaye, J.; Lalling

Proc. Natl. Acad. Sci. U.S.A. 87, 6363-6367, 1990

A;Title: Genes for Plasmodium falciparum surface antigens cloned by expression in COS ce

A;Reference number: A36018; MUID:90349616

A;Accession: A36018

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-287 <ELL>

A;Cross-references: GB:M28890; NID:gl60406; PID:gl60407

R;Smythe, J.A.; Peterson, M.G.; Coppel, R.L.; Saul, A.J.; Kemp, D.J.; Anders, R.F.

Mol. Biochem. Parasitol. 39, 227-234, 1990

A;Title: Structural diversity in the 45-kilodalton merozoite surface antigen of Plasmodi

A;Reference number: A44950; MUID:90205972

A;Accession: B44950

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-186,'R',188-287 <SMY>

A;Cross-references: GB:M28892; NID:gl60488; PID:gl60489

R;Fandeur, T.; Bonnefoy, S.; Mercereau-Pujalon, O.

Mol. Biochem. Parasitol. 47, 167-178, 1991

A;Title: In vivo and in vitro derived Palo Alto lines of Plasmodium falciparum are ge

A;Reference number: A45613; MUID:92049549

A;Accession: A45613

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: nucleic acid

A;Residues: 28,'E',30-186,'R',188-230,'IH' <FAN>

A;Experimental source: Uganda Palo Alto strain, merozoite

A;Note: sequence extracted from NCBI backbone (NCBIP:65035)

C;Superfamily: Epstein-Barr virus nuclear antigen

C;Keywords: membrane protein; surface antigen

Query Match 100.0%; Score 41; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SNTFINNA 8
| | | | | | | |
Db 27 SNTFINNA 34

RESULT 10

A39112

merozoite 45K surface antigen precursor - malaria parasite (Plasmodium falciparum) (1

C;Species: Plasmodium falciparum

C;Date: 30-Aug-1991 #sequence_revision 30-Aug-1991 #text_change 17-Nov-2000

C;Accession: A39112

R;Smythe, J.A.; Coppel, R.L.; Day, K.P.; Martin, R.K.; Oduola, A.M.J.; Kemp, D.J.; An

Proc. Natl. Acad. Sci. U.S.A. 88, 1751-1755, 1991

A;Title: Structural diversity in the Plasmodium falciparum merozoite surface antigen

A;Reference number: A39112; MUID:91156685

A;Accession: A39112

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-300 <SMY>

A;Cross-references: GB:M59765

C;Superfamily: Epstein-Barr virus nuclear antigen

C;Keywords: surface antigen

Query Match 100.0%; Score 41; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SNTFINNA 8
| | | | | | | |
Db 27 SNTFINNA 34

RESULT 11

A39615

merozoite 45K surface antigen precursor (clone T9-96) - malaria parasite (Plasmodium

C;Species: Plasmodium falciparum

C;Date: 13-Sep-1991 #sequence_revision 13-Sep-1991 #text_change 17-Nov-2000

C;Accession: A39615; S13802

R;Fenton, B.; Clark, J.T.; Khan, C.M.A.; Robinson, J.V.; Walliker, D.; Ridley, R.; Sc

Mol. Cell. Biol. 11, 963-971, 1991

A;Title: Structural and antigenic polymorphism of the 35- to 48-kilodalton merozoite

A;Reference number: A39615; MUID:91117264

A;Accession: A39615

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-302 <FEN>

A;Cross-references: EMBL:X53832

C;Superfamily: Epstein-Barr virus nuclear antigen

C;Keywords: surface antigen

Query Match 100.0%; Score 41; DB 2; Length 302;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
 |||||
 Db 27 SNTFINNA 34

RESULT 12

B39112
 merozoite 45K surface antigen precursor - malaria parasite (Plasmodium falciparum) (isol
 C:Species: Plasmodium falciparum
 C:Date: 30-Aug-1991 #sequence_revision 30-Aug-1991 #text_change 17-Nov-2000
 C:Accession: B39112
 R:Smythe, J.A.; Coppel, R.L.; Day, K.P.; Martin, R.K.; Oduola, A.M.J.; Kemp, D.J.; Ander
 Proc. Natl. Acad. Sci. U.S.A. 88, 1751-1755, 1991
 A:Title: Structural diversity in the Plasmodium falciparum merozoite surface antigen 2.
 A:Reference number: A39112; MUID:91156685
 A:Accession: B39112
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-347 <SMY>
 A:Cross-references: GB:M59767
 C:Superfamily: Epstein-Barr virus nuclear antigen
 C:Keywords: surface antigen

Query Match 100.0%; Score 41; DB 2; Length 347;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
 |||||
 Db 27 SNTFINNA 34

RESULT 13

S78382
 ribosomal protein S4, plastid - beechdrops plastid
 C:Species: plastid Epifagus virginiana (beechdrops)
 C:Date: 25-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 13-Aug-1999
 C:Accession: S78382
 R:Wolfe, K.H.; Morden, C.W.; Ems, S.C.; Palmer, J.D.
 J. Mol. Evol. 35, 304-317, 1992
 A:Title: Rapid evolution of the plastid translational apparatus in a nonphotosynthetic P
 A:Reference number: S78378; MUID:93021155
 A:Accession: S78382
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-202 <WOL>
 A:Cross-references: EMBL:M81884; NID:g336917; PIDN:AA65853.1; PID:g336927
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1992
 C:Genetics:
 A:Gene: rps4
 A:Genome: plastid
 C:Superfamily: Escherichia coli ribosomal protein S4
 C:Keywords: plastid; protein biosynthesis; ribosome

Query Match 90.2%; Score 37; DB 2; Length 202;
 Best Local Similarity 100.0%; Pred. No. 4.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINN 7
 |||||
 Db 144 SNTFINN 150

RESULT 14

NMIVEA
 exo-alpha-sialidase (EC 3.2.1.18) - influenza A virus (strain A/Cor/16/74[N7])
 N:Alternate names: neuraminidase

C:Species: Influenza A virus
 C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
 C:Accession: A25615
 R:Dale, B.; Brown, R.; Miller, J.; White, R.T.; Air, G.M.; Cordell, B.
 Virology 155, 460-468, 1986
 A:Title: Nucleotide and deduced amino acid sequence of the influenza neuraminidase ge
 A:Reference number: A94349; MUID:87071664
 A:Accession: A25615
 A:Molecule type: mRNA
 A:Residues: 1-469 <DAL>

A:Cross-references: GB:M14916; NID:g323979; PIDN:AAA43093.1; PID:g323980
 A:Note: this enzyme exists as a tetramer in the viral membrane; the amino-terminal hy
 C:Comment: This enzyme catalyzes the cleavage of the terminal sialic acid (N-acetylme
 ing the carbohydrate residues from the viral envelope.
 C:Genetics:
 A:Map position: segment 6
 C:Superfamily: Influenza virus exo-alpha-sialidase
 C:Keywords: glycoprotein; glycosidase; hydrolase; transmembrane protein
 F:8-24/Donah: transmembrane #status predicted <TMW>
 F:25-88/Region: hypervariable stalk
 F:89-469/Region: head of exo-alpha-sialidase
 F:28,32,46,55,66,85,143,198,232,399/Binding site: carbohydrate (Asn) (covalent) #stat

Query Match 82.9%; Score 34; DB 1; Length 469;
 Best Local Similarity 85.7%; Pred. No. 37;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 NTFINNA 8
 ||:||||
 Db 61 NTFINNA 67

RESULT 15

F56653
 virion protein - human herpesvirus 6 (strain UI102) (fragment)
 C:Species: human herpesvirus 6
 C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 02-Jun-2000
 C:Accession: F56653
 R:Gompels, U.A.; Carss, A.L.; Sun, N.; Arrand, J.R.
 DNA Seq. 3, 25-39, 1992
 A:Title: Infectivity determinants encoded in a conserved gene block of human herpesvi
 A:Reference number: A56653; MUID:93091236
 A:Accession: F56653
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-378 <GOM>
 A:Cross-references: GB:X64320; GB:S50422; GB:S58505; GB:S58506; NID:g296190; PIDN:CAA
 A:Experimental source: strain UI102
 A:Note: sequence extracted from NCBI backbone (NCBIN:120533, NCBIP:120551)
 C:Superfamily: varicella-zoster virus gene 34 protein

Query Match 80.5%; Score 33; DB 2; Length 378;
 Best Local Similarity 85.7%; Pred. No. 46;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NTFINNA 8
 ||| |||
 Db 141 NTFINNA 147

Search completed: January 29, 2002, 10:26:42
 Job time: 653 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:13:44 ; Search time 80.65 seconds
(without alignments)
3.637 Million cell updates/sec

Title: US-09-763-397A-14
Perfect score: 41
Sequence: 1 SNTFINNA 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	41	100.0	262	1 MSA2_PLAF6	Q93317 plasmodium
2	41	100.0	264	1 MSA2_PLAF7	P19599 plasmodium
3	41	100.0	264	1 MSA2_PLAF7	P50499 plasmodium
4	41	100.0	272	1 MSA2_PLAF7	P50498 plasmodium
5	41	100.0	274	1 MSA2_PLAF6	P50497 plasmodium
6	41	100.0	276	1 MSA2_PLAF8	Q93320 plasmodium
7	41	100.0	280	1 MSA2_PLAF6	Q93643 plasmodium
8	41	100.0	281	1 MSA2_PLAFH	Q93319 plasmodium
9	41	100.0	286	1 MSA2_PLAF1	P50496 plasmodium
10	41	100.0	287	1 MSA2_PLAF1	P19260 plasmodium
11	41	100.0	300	1 MSA2_PLAF1	Q93644 plasmodium
12	41	100.0	300	1 MSA2_PLAF2	Q93645 plasmodium
13	41	100.0	302	1 MSA2_PLAF9	Q93994 plasmodium
14	41	100.0	347	1 MSA2_PLAF2	Q93646 plasmodium
15	37	90.2	202	1 MSA2_PLAF1	P30056 epifluenza vi
16	34	82.9	469	1 NRAM_IHCO	P08327 influenza a
17	33	80.5	378	1 UL25_HSV6G	P52536 human herpe
18	33	80.5	436	1 YIN5_YEAST	P40463 saccharomyc
19	33	80.5	555	1 UL25_HSV6U	P52387 human herpe
20	33	80.5	555	1 UL25_HSV6Z	P52537 human herpe
21	32	78.0	320	1 O7C1_HUMAN	O76099 homo sapien
22	32	78.0	320	1 O7C2_HUMAN	O60412 homo sapien
23	32	78.0	351	1 COA2_POVBA	P14997 polyomaviru
24	32	78.0	351	1 COA2_POVBA	P30094 polyomaviru
25	32	78.0	614	1 GAAL_YEAST	P39012 saccharomyc
26	32	78.0	644	1 HM38_CABEL	UL9720 caenorhabdi
27	31	75.6	110	1 RR18_PEA	P49169 pisum sativ
28	31	75.6	380	1 IPR7_PLAF7	O77392 plasmodium
29	31	75.6	519	1 LAC2_TRAVE	Q12718 trameetes ve
30	31	75.6	519	1 LAC2_TRAVI	Q99046 trameetes vi
31	31	75.6	558	1 YNT3_YEAST	P53870 saccharomyc
32	31	75.6	567	1 NU2M_HANWI	P48906 hansenula w
33	31	75.6	573	1 YEF8_YEAST	P40095 saccharomyc

34	31	75.6	700	1 CAN2_CHICK	Q92178 gallus gall
35	31	75.6	2469	1 TEGU_HSVSA	Q01056 herpesvirus
36	30	73.2	106	1 YNE3_YEAST	P53957 saccharomyc
37	30	73.2	144	1 GIB_APLKU	P02211 aplysia kur
38	30	73.2	146	1 GIB_APLLI	P02210 aplysia lim
39	30	73.2	214	1 YOGG_BACST	P28754 bacillus st
40	30	73.2	277	1 MALD_STRPN	Q04699 streptococc
41	30	73.2	309	1 OLF4_CANFA	Q95157 canis famli
42	30	73.2	340	1 VB17_VACCC	P21075 vaccinia vi
43	30	73.2	340	1 VB17_VACCV	Q01221 vaccinia vi
44	30	73.2	340	1 VB17_VARV	P33878 variola vir
45	30	73.2	430	1 YOGG_BACSU	P54487 bacillus su

ALIGNMENTS

RESULT 1
MSA2_PLAF6
ID MSA2_PLAF6 STANDARD; PRT: 262 AA.
AC Q9317;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MEROZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2) (ALLELIC FORM 1).
GN MSA2.
OS Plasmodium falciparum (isolate Camp / Malaysia).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5835;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91218803; PubMed=2090943;
RA Thomas A.W., Carr D.A., Carter J.M., Lyon J.A.;
RT "Sequence comparison of allelic forms of the Plasmodium falciparum merozoite surface antigen MSA2.";
RL Mol. Biochem. Parasitol. 43:211-220(1990).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE ERYTHROCYTE.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (POTENTIAL).
CC -!- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
EMBL: M60186; AAA29687.1; -
InterPro: IPR001136; MSA_2.
Pfam: PF00985; MSA_2; 1.
Malaria: Membrane; Glycoprotein; Antigen; Signal; Repeat;
GPI-anchor; Merozoite.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 238 MEROZOITE SURFACE ANTIGEN 2.
FT PROPEP 239 262 HYDROPHOBIC, REMOVED DURING MATURATION (BY SIMILARITY).
FT DOMAIN 44 188 POLYMORPHIC REGION.
FT CARBOHYD 91 98 POLY-THR.
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 211 211 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 262 AA; 27374 MW; 72E0B2A315E9D154 CRC64;

Query Match 100.0%; Score 41; DB 1; Length 262;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

...

DE SURFACE ANTIGEN).
 OS MSA2.
 GN Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90205972; PubMed=21811307;
 RA Smythe J.A., Peterson M.G., Coppel R.L., Saul A.J., Kemp D.J.,
 RA Anders R.F.;
 RT "Structural diversity in the 45-kilodalton merozoite surface antigen
 of Plasmodium falciparum.";
 RL Mol. Biochem. Parasitol. 39:227-234(1990).
 CC -!- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE
 CC ERYTHROCYTE.
 CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
 CC (POTENTIAL).
 CC -!- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M28891; AAA29686.1; -;
 DR InterPro: IPR001136; MSA_2.
 DR Pfam: PF00985; MSA_2; 1.
 KW Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
 KW GPI-anchor; Merozoite.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 248 MEROZOITE SURFACE ANTIGEN 2.
 FT PROPEP 249 272 HYDROPHOBIC, REMOVED DURING MATURATION
 FT (BY SIMILARITY).
 FT DOMAIN 44 198 POLY-THR.
 FT DOMAIN 95 108 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 22 22 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 36 36 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 149 149 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 221 221 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 245 245 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 272 AA; 27971 MW; 9D9CF223BF2B483D CRC64;
 Query Match 100.0%; Score 41; DB 1; Length 272;
 Best Local Similarity 100.0%; Pred. No. 0.22;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNTFINNA 8
 Db [11111111]
 Db 27 SNTFINNA 34
 RESULT 5
 MSA2_PLAF6 STANDARD; PRT; 274 AA.
 ID MSA2_PLAF6
 AC P50497;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE MEROZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2).
 GN MSA2.
 OS Plasmodium falciparum (isolate kf1916).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=57269;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92178286; PubMed=1542312;
 RA Marshall V.M., Coppel R.L., Anders R.F., Kemp D.J.;
 RT "Two novel alleles within subfamilies of the merozoite surface

RT antigen 2 (MSA-2) of Plasmodium falciparum.";
 RL Mol. Biochem. Parasitol. 50:181-184(1992).
 CC -!- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE
 CC ERYTHROCYTE.
 CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
 CC (POTENTIAL).
 CC -!- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M73810; AAA29698.1; -;
 DR InterPro: IPR001136; MSA_2.
 DR Pfam: PF00985; MSA_2; 1.
 KW Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
 KW GPI-anchor; Merozoite.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 250 MEROZOITE SURFACE ANTIGEN 2.
 FT PROPEP 251 274 HYDROPHOBIC, REMOVED DURING MATURATION
 FT (BY SIMILARITY).
 FT DOMAIN 44 200 POLY-THR.
 FT DOMAIN 97 110 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 22 22 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 36 36 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 223 223 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 248 248 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 274 AA; 28367 MW; CEA832D766F743A2 CRC64;
 Query Match 100.0%; Score 41; DB 1; Length 274;
 Best Local Similarity 100.0%; Pred. No. 0.22;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNTFINNA 8
 Db [11111111]
 Db 27 SNTFINNA 34
 RESULT 6
 MSA2_PLAF8 STANDARD; PRT; 276 AA.
 ID MSA2_PLAF8
 AC Q99320;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE MEROZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2) (ALLELIC FORM 4).
 GN MSA2.
 OS Plasmodium falciparum (isolate 7G8).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=57266;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91218803; PubMed=2090943;
 RA Thomas A.W., Carr D.A., Carter J.M., Lyon J.A.;
 RT "Sequence comparison of allelic forms of the Plasmodium falciparum
 merozoite surface antigen MSA2.";
 RL Mol. Biochem. Parasitol. 43:211-220(1990).
 CC -!- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE
 CC ERYTHROCYTE.
 CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
 CC (POTENTIAL).
 CC -!- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M60190; AAA29690.1; -
InterPro; IPR001136; MSA_2.
Pfam; PF00985; MSA_2; 1.
Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
GPI-anchor; Merozoite.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 252 MEROZOITE SURFACE ANTIGEN 2.
FT PROPEP 253 276 HYDROPHOBIC, REMOVED DURING MATURATION
(BY SIMILARITY).
FT DOMAIN 44 202 POLYMORPHIC REGION.
FT CARBOHYD 105 112 POLY-THR.
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 225 225 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 276 AA; 28172 MW; 85FA62A70400DEB2 CRC64;

Query Match 100.0%; Score 41; DB 1; Length 276;
Best Local Similarity 100.0%; Pred. No. 0.23; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNTFINNA 8
DB 27 SNTFINNA 34

RESULT 7
MSA2_PLAFK STANDARD; PRT; 280 AA.
AC O03643;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MEROZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2).
GN MSA2.
OS Plasmodium falciparum (isolate K1 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5839;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91156685; PubMed=2000383;
RA Smythe J.A., Coppel R.L., Day K.P., Martin R.K., Oduola A.M.J.,
RA Kemp D.J., Anders R.F.;
RT "Structural diversity in the Plasmodium falciparum merozoite surface antigen 2";
RL Proc. Natl. Acad. Sci. U.S.A. 88:1751-1755(1991).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE ERYTHROCYTE.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (POTENTIAL).
CC -!- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M59766; AAA29693.1; -
InterPro; IPR001136; MSA_2.
Pfam; PF00985; MSA_2; 1.
Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
GPI-anchor; Merozoite.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 256 MEROZOITE SURFACE ANTIGEN 2.

FT PROPEP 257 280 HYDROPHOBIC, REMOVED DURING MATURATION
(BY SIMILARITY).
FT DOMAIN 44 206 POLYMORPHIC REGION.
FT CARBOHYD 160 163 POLY-SER.
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 229 229 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 280 AA; 29542 MW; 9395D8AFB6B473EF CRC64;

Query Match 100.0%; Score 41; DB 1; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.23; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNTFINNA 8
DB 27 SNTFINNA 34

RESULT 8
MSA2_PLAFH STANDARD; PRT; 281 AA.
AC Q99319.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MEROZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2) (ALLELIC FORM 3).
GN MSA2.
OS Plasmodium falciparum (isolate thtn / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=70151;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91218803; PubMed=2090943;
RA Thomas A.W., Carr D.A., Carter J.M., Lyon J.A.;
RT "Sequence comparison of allelic forms of the Plasmodium falciparum merozoite surface antigen MSA2";
RL Mol. Biochem. Parasitol. 43:211-220(1990).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE ERYTHROCYTE.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (POTENTIAL).
CC -!- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M60189; AAA29689.1; -
InterPro; IPR001136; MSA_2.
Pfam; PF00985; MSA_2; 1.
Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
GPI-anchor; Merozoite.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 257 MEROZOITE SURFACE ANTIGEN 2.
FT PROPEP 258 281 HYDROPHOBIC, REMOVED DURING MATURATION
(BY SIMILARITY).
FT DOMAIN 44 207 POLYMORPHIC REGION.
FT CARBOHYD 111 118 POLY-THR.
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 281 AA; 28892 MW; 50598AA42D64CCBC CRC64;

Query Match 100.0%; Score 41; DB 1; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.23; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

QY 1 SNTFINNA 8
| | | | | | | |
DB 27 SNTFINNA 34

RESULT 9
MSA2_PLAF1 STANDARD; PRT; 286 AA.
AC P50496;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MERZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2).
GN MSA2.
OS Plasmodium falciparum (isolate 311).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57265;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92178286; PubMed=1542312;
RA Marshall V.M., Coppel R.L., Anders R.F., Kemp D.J.;
RT "Two novel alleles within subfamilies of the merozoite surface antigen 2 (MSA-2) of Plasmodium falciparum."
RL Mol. Biochem. Parasitol. 50:181-184(1992).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE MERZOITE ATTACHMENT TO THE ERYTHROCYTE.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (POTENTIAL).
CC -1- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
CC EMBL: M73809; AAA29697.1; -;
DR InterPro: IPR001136; MSA_2.
DR Pfam: PF00985; MSA_2; 1.
KW Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
FT GPI-anchor; Merozoite.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 262 MERZOITE SURFACE ANTIGEN 2.
FT PROPEP 263 286 HYDROPHOBIC, REMOVED DURING MATURATION (BY SIMILARITY).
FT DOMAIN 44 212 POLYMORPHIC REGION.
FT DOMAIN 115 122 POLY-THR.
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 286 AA; 28844 MW; D1F4947CE68D5805 CRC64;

Query Match 100.0%; Score 41; DB 1; Length 286;
Best Local Similarity 100.0%; Pred. No. 0.24; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

QY 1 SNTFINNA 8
| | | | | | | |
DB 27 SNTFINNA 34

RESULT 10
MSA2_PLAFG STANDARD; PRT; 287 AA.
AC P19260;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MERZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2) (ALLELIC FORM 2) (MEMBRANE PROTEIN PF7).
GN MSA2.
OS Plasmodium falciparum (isolate FCR-3 / Gambia).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5838;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90349616; PubMed=1696728;
RA Elliott J.F., Albrecht G.R., Gilladoga A., Handunnetti S.M., Neequaye J., Lallinger G., Minjas J.N., Howard R.J.;
RT "Genes for Plasmodium falciparum surface antigens cloned by expression in COS cells."
RL Proc. Natl. Acad. Sci. U.S.A. 87:6363-6367(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91117264; PubMed=1990294;
RA Fenton B., Clark J.T., Khan C.M.A., Robinson J.V., Walliker D., Ridley R., Scaife J.G., McBride J.S.;
RT "Structural and antigenic polymorphism of the 35- to 48-kilodalton merozoite surface antigen (MSA-2) of the malaria parasite Plasmodium falciparum."
RL Mol. Cell. Biol. 11:963-971(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91218803; PubMed=2090943;
RA Thomas A.W., Carr D.A., Carter J.M., Lyon J.A.;
RT "Sequence comparison of allelic forms of the Plasmodium falciparum merozoite surface antigen MSA2."
RL Mol. Biochem. Parasitol. 43:211-220(1990).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE MERZOITE ATTACHMENT TO THE ERYTHROCYTE.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (POTENTIAL).
CC -1- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
CC EMBL: M28890; AAA29650.1; -;
DR EMBL: X53832; CAA37829.1; -;
DR EMBL: M60188; AAA29688.1; -;
DR PIR: A36018; A36018.
DR PIR: B39615; B39615.
DR InterPro: IPR001136; MSA_2.
DR Pfam: PF00985; MSA_2; 1.
KW Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
FT GPI-anchor; Merozoite.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 263 MERZOITE SURFACE ANTIGEN 2.
FT PROPEP 264 287 HYDROPHOBIC, REMOVED DURING MATURATION (BY SIMILARITY).
FT DOMAIN 44 213 POLYMORPHIC REGION.
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 287 AA; 28555 MW; 3968B90DAA917AF8 CRC64;

RESULT 12					
MSA2_PLAF2					
ID	MSA2_PLAF2	STANDARD;	PRT;	300 AA.	
AC	Q03645;				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	01-OCT-1996 (Rel. 34, Last annotation update)				
DE	MEROZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2).				
GN	MSA2.				
OS	Plasmodium falciparum (isolate mad71 / Papua New Guinea).				
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.				
OX	NCBI_TaxID=70154;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=91156685; PubMed=2000383;				
RA	Smythe J.A., Coppel R.L., Day K.P., Martin R.K., Oduola A.M.J.,				
RA	Kemp D.J., Anders R.F.;				
RT	"Structural diversity in the Plasmodium falciparum merozoite surface				
RT	antigen 2.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 88:1751-1755(1991).				
CC	-!- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE				
CC	ERYTHROCYTE.				
CC	-!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR				
CC	(POTENTIAL).				
CC	-!- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	EMBL; M59768; AAA29696.1; -				
DR	PIR; A39112; A39112.				
DR	InterPro; IPR001136; MSA_2.				
DR	Pfam; PF00985; MSA_2; 1.				
KW	Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;				
KW	GPI-anchor; Merozoite.				
FT	SIGNAL 1 20				
FT	CHAIN 21 276				
FT	PROPEP 277 300				
FT	POTENTIAL.				
FT	MEROZOITE SURFACE ANTIGEN 2.				
FT	HYDROPHOBIC, REMOVED DURING MATURATION				
FT	(BY SIMILARITY).				
FT	POLYMORPHIC REGION.				
FT	POLY-THR.				
FT	DOMAIN 44 226				
FT	DOMAIN 129 136				
FT	CARBOHYD 22 22				
FT	CARBOHYD 36 36				
FT	CARBOHYD 177 177				
FT	CARBOHYD 249 249				
FT	CARBOHYD 273 273				
FT	CARBOHYD 274 274				
FT	SEQUENCE 300 AA; 30101 MW; E4116107747AA10D CRC64;				
Qy	1 SNTFINNA 8				
Db	27 SNTFINNA 34				
Query Match					
Best Local Similarity		100.0%;	Score 41;	DB 1;	Length 300;
Matches		8;	Conservative	0;	Mismatches
				0;	Indels
				0;	Gaps
RESULT 13					
MSA2_PLAF9					
ID	MSA2_PLAF9	STANDARD;	PRT;	302 AA.	
AC	Q03994;				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	01-OCT-1996 (Rel. 34, Last annotation update)				
DE	MEROZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2).				

Query Match		100.0%;	Score 41;	DB 1;	Length 287;
Best Local Similarity		100.0%;	Pred. No. 0.24;		
Matches		8;	Conservative	0;	Mismatches
				0;	Indels
				0;	Gaps
Y	1 SNTFINNA 8				
b	27 SNTFINNA 34				
RESULT 11					
MSA2_PLAFI	MSA2_PLAFI	STANDARD;	PRT;	300 AA.	
ID	Q03644;				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	15-DEC-1998 (Rel. 37, Last annotation update)				
DE	MEZOZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2).				
GN	MSA2.				
OS	Plasmodium falciparum (isolate imr143).				
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.				
OX	NCBI_TaxID=57268;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=91156685; PubMed=2000383;				
RA	Smythe J.A., Coppel R.L., Day K.P., Martin R.K., Oduola A.M.J.,				
RA	Kemp D.J., Anders R.F.;				
RT	"Structural diversity in the Plasmodium falciparum merozoite surface				
RT	antigen 2.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 88:1751-1755(1991).				
CC	-!- FUNCTION: MAY PLAY A ROLE IN THE MEZOZOITE ATTACHMENT TO THE				
CC	ERYTHROCYTE.				
CC	-!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR				
CC	(POTENTIAL).				
CC	-!- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	EMBL; M59767; AAA29695.1; -				
DR	InterPro; IPR001136; MSA_2.				
DR	Pfam; PF00985; MSA_2; 1.				
KW	Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;				
KW	GPI-anchor; Merozoite.				
FT	SIGNAL 1 20				
FT	CHAIN 21 276				
FT	PROPEP 277 300				
FT	POTENTIAL.				
FT	MEZOZOITE SURFACE ANTIGEN 2.				
FT	HYDROPHOBIC, REMOVED DURING MATURATION				
FT	(BY SIMILARITY).				
FT	POLYMORPHIC REGION.				
FT	POLY-THR.				
FT	DOMAIN 44 226				
FT	DOMAIN 129 136				
FT	CARBOHYD 22 22				
FT	CARBOHYD 36 36				
FT	CARBOHYD 177 177				
FT	CARBOHYD 249 249				
FT	CARBOHYD 273 273				
FT	CARBOHYD 274 274				
FT	SEQUENCE 300 AA; 30131 MW; A01E17D36075D7D6 CRC64;				
QY	1 SNTFINNA 8				
Db	27 SNTFINNA 34				
Query Match		100.0%;	Score 41;	DB 1;	Length 300;
Best Local Similarity		100.0%;	Pred. No. 0.25;		
Matches		8;	Conservative	0;	Mismatches
				0;	Indels
				0;	Gaps

GN MSA2.
 OS Plasmodium falciparum (isolate tak 9).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=57276;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91117264; PubMed=1990294;
 RA Fenton B., Clark J.T., Khan C.M.A., Robinson J.V., Walliker D.,
 RT "Structural and antigenic polymorphism of the 35- to 48-kilodalton
 merozoite surface antigen (MSA-2) of the malaria parasite Plasmodium
 falciparum.";
 RL Mol. Cell. Biol. 11:963-971(1991).
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE
 CC ERYTHROCYTE.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
 CC (POTENTIAL).
 CC -1- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X53833; CAA37830.1; -;
 DR PIR; A39615; A39615.
 DR InterPro; IPR001136; MSA_2.
 DR Pfam; PF00985; MSA_2; 1.
 KW Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
 KW GPI-anchor; Merozoite.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 278 MEROZOITE SURFACE ANTIGEN 2.
 FT PROPEP 279 302 HYDROPHOBIC, REMOVED DURING MATURATION
 FT (BY SIMILARITY).
 FT DOMAIN 44 228 POLYMORPHIC REGION.
 FT DOMAIN 131 138 POLY-THR.
 FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 302 AA; 30259 MW; 4E0A7EB08227CF66 CRC64;

Query Match 100.0%; Score 41; DB 1; Length 302;
 Best Local Similarity 100.0%; Pred. No. 0.25;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
 Db 27 SNTFINNA 34
 |||||

RESULT 14
 MSA2_PLAF2
 ID MSA2_PLAF2 STANDARD; PRT; 347 AA.
 AC Q03646;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE MEROZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2).
 GN MSA2.
 OS Plasmodium falciparum (isolate Nig32 / Nigeria).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=70150;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91156685; PubMed=2000383;
 RA Smythe J.A., Coppel R.L., Day K.P., Martin R.K., Oduola A.M.J.,

RA Kemp D.J., Anders R.F.;
 RT "Structural diversity in the Plasmodium falciparum merozoite surface
 antigen 2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:1751-1755(1991).
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE
 CC ERYTHROCYTE.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
 CC (POTENTIAL).
 CC -1- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M59765; AAA28691.1; -;
 DR PIR; B39112; B39112.
 DR InterPro; IPR001136; MSA_2.
 DR Pfam; PF00985; MSA_2; 1.
 KW Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
 KW GPI-anchor; Merozoite.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 323 MEROZOITE SURFACE ANTIGEN 2.
 FT PROPEP 324 347 HYDROPHOBIC, REMOVED DURING MATURATION
 FT (BY SIMILARITY).
 FT DOMAIN 44 273 POLYMORPHIC REGION.
 FT DOMAIN 177 184 POLY-THR.
 FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 224 224 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 347 AA; 33786 MW; ABCF24BB560BF537 CRC64;

Query Match 100.0%; Score 41; DB 1; Length 347;
 Best Local Similarity 100.0%; Pred. No. 0.29;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
 Db 27 SNTFINNA 34
 |||||

RESULT 15
 RR4_EPVI
 ID RR4_EPVI STANDARD; PRT; 202 AA.
 AC P30056;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S4.
 GN RPS4.
 OS Epifagus virginiana (Beechdrops).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Lamiales; Orobanchaceae; Epifagus.
 OX NCBI_TaxID=4177;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93066301; PubMed=1332054;
 RA Wolfe K.H., Morden C.W., Palmer J.D.;
 RT "Function and evolution of a minimal plastid genome from a
 nonphotosynthetic parasitic plant.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:10648-10652(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93021155; PubMed=1404416;

Wolfe K.H., Morden C.W., Ems S.C., Palmer J.D.;
 "Rapid evolution of the plastid translational apparatus in a
 nonphotosynthetic plant: loss or accelerated sequence evolution of
 tRNA and ribosomal protein genes.";
 J. Mol. Evol. 35:304-317(1992).
 -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA
 (BY SIMILARITY).
 -!- SIMILARITY: CONTAINS 1 S4 RNA-BINDING DOMAIN.
 -!- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; M81884; AAA65853.1; -;
 Mendel; 4196; EPiVi; rps4.1.
 InterPro: IPR001912; Ribosomal_S4.
 InterPro: IPR002942; S4.
 Pfam; PF00163; Ribosomal_S4; 1.
 Pfam; PF01479; S4; 1.
 SMART; SW00363; S4; 1.
 PROSITE; PS00632; RIBOSOMAL_S4; 1.
 Ribosomal protein; rRNA-binding; Chloroplast.
 DOMAIN 89 136 RNA-BINDING (S4_TYPE).
 FT SEQUENCE 202 AA; 23824 MW; 78857286516BA691 CRC64;
 SQ

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:12:13 ; Search time 285.36 Seconds
(without alignments)
4.101 Million cell updates/sec

Title: US-09-763-397A-14
Perfect score: 41
Sequence: 1 SNTFINNA 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 14627329 residues 473505
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_17.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_prodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	41	100.0	108	5	Q9BJP0		Q9BJP0 plasmodium
2	41	100.0	111	5	Q94666		Q94666 plasmodium
3	41	100.0	113	5	Q9BJN8		Q9BJN8 plasmodium
4	41	100.0	116	5	Q9GQ26		Q9GQ26 plasmodium
5	41	100.0	117	5	Q9GQX2		Q9GQX2 plasmodium
6	41	100.0	117	5	Q9BJP9		Q9BJP9 plasmodium
7	41	100.0	118	5	Q9BJP2		Q9BJP2 plasmodium
8	41	100.0	119	5	Q9GQ20		Q9GQ20 plasmodium
9	41	100.0	119	5	Q9BJP7		Q9BJP7 plasmodium
10	41	100.0	120	5	Q9GQY0		Q9GQY0 plasmodium
11	41	100.0	122	5	Q9BJN7		Q9BJN7 plasmodium
12	41	100.0	125	5	Q9GQY9		Q9GQY9 plasmodium
13	41	100.0	128	5	Q9GQ24		Q9GQ24 plasmodium
14	41	100.0	129	5	Q9GQ25		Q9GQ25 plasmodium
15	41	100.0	131	5	Q94664		Q94664 plasmodium
16	41	100.0	131	5	Q94672		Q94672 plasmodium
17	41	100.0	131	5	Q9GQ29		Q9GQ29 plasmodium
18	41	100.0	132	5	Q9GR02		Q9GR02 plasmodium
19	41	100.0	135	5	O00785		O00785 plasmodium

20	41	100.0	135	5	Q9GR00	Q9gr00 plasmodium
21	41	100.0	135	5	Q9BJP4	Q9bjp4 plasmodium
22	41	100.0	136	5	Q9GQ28	Q9gq28 plasmodium
23	41	100.0	136	5	Q9BJR5	Q9bjr5 plasmodium
24	41	100.0	136	5	Q9BJP3	Q9bjp3 plasmodium
25	41	100.0	136	5	Q9BJP1	Q9bjp1 plasmodium
26	41	100.0	136	5	Q9BJN2	Q9bjn2 plasmodium
27	41	100.0	137	5	Q9GQX7	Q9gqx7 plasmodium
28	41	100.0	138	5	Q9GQY8	Q9gqy8 plasmodium
29	41	100.0	139	5	Q9GQ22	Q9gq22 plasmodium
30	41	100.0	139	5	Q9GQX9	Q9gqx9 plasmodium
31	41	100.0	141	5	Q9BJR9	Q9bjr9 plasmodium
32	41	100.0	141	5	Q9BJP8	Q9bjp8 plasmodium
33	41	100.0	144	5	Q9GQY4	Q9gqy4 plasmodium
34	41	100.0	145	5	Q9GR03	Q9gr03 plasmodium
35	41	100.0	149	5	Q94670	Q94670 plasmodium
36	41	100.0	150	5	Q9GQY6	Q9gqy6 plasmodium
37	41	100.0	150	5	Q9GQX3	Q9gqx3 plasmodium
38	41	100.0	150	5	Q9BJQ7	Q9bjq7 plasmodium
39	41	100.0	150	5	Q9BJQ6	Q9bjq6 plasmodium
40	41	100.0	151	5	Q9BJN5	Q9bjn5 plasmodium
41	41	100.0	152	5	Q9GR04	Q9gr04 plasmodium
42	41	100.0	152	5	Q9BJS5	Q9bjs5 plasmodium
43	41	100.0	152	5	Q9BJS1	Q9bjs1 plasmodium
44	41	100.0	153	5	Q94667	Q94667 plasmodium
45	41	100.0	153	5	Q9GQY3	Q9gqy3 plasmodium

ALIGNMENTS

RESULT 1

Q9BJP0 PRELIMINARY; PRT; 108 AA.

AC Q9BJP0; 01-JUN-2001 (TREMblrel. 17, Created)

DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)

DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)

DE MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5833;

RN [1]

RP SEQUENCE FROM N.A.

RA Eisen D.P., Clonnan N., Baddeley A., Eri R., Saul A.;

RT "Antigenic drift and immune selection acting on merozoite surface protein 1-19 and merozoite surface protein 2 in independent field isolates of Plasmodium falciparum.";

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF329573; AAK19392.1; -

FT NON_TER 1

FT NON_TER 108

SQ SEQUENCE 108 AA; 10897 MW; C317035F0D1A46C8 CRC64;

Query Match 100.0%; Score 41; DB 5; Length 108;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
| | | | | | | |
Db 10 SNTFINNA 17

RESULT 2

Q94666 PRELIMINARY; PRT; 111 AA.

ID Q94666; 01-FEB-1997 (TREMblrel. 02, Created)

AC Q94666; 01-FEB-1997 (TREMblrel. 02, Last sequence update)

DT 01-FEB-1997 (TREMblrel. 17, Last annotation update)

DE MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).

GN MSP-2.

OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VN9; Wang L., Billman-Jacobe H., Hanh Nhan D., Richie T.L.,
RA Weisman S., Coppel R.L.;
RA "Antibody responses to infections with strains of Plasmodium
RT falciparum expressing diverse forms of Merozoite Surface Protein 2.";
RL Infect. Immun. 0:0-0(2001).
DR EMBL; AF104692; AAG47600.1; -
DR InterPro; IPR001136; MSA_2.
DR Pfam; PF00985; MSA_2; 1.
FT NON_TER 1
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 11854 MW; 6BF44658DD9497BA CRC64;

Query Match 100.0%; Score 41; DB 5; Length 116;
Best Local Similarity 100.0%; Pred. No. 0.46; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

QY 1 SNTFINNA 8
Db |11111111|
Db 10 SNTFINNA 17

RESULT 5
Q9GQX2 PRELIMINARY; PRT; 117 AA.
AC Q9GQX2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE MEROZOITE SURFACE PROTEIN (FRAGMENT).
GN MSP2.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VN33;
RA Weisman S., Wang L., Billman-Jacobe H., Hanh Nhan D., Richie T.L.,
RA Coppel R.L.;
RT "Antibody responses to infections with strains of Plasmodium
RT falciparum expressing diverse forms of Merozoite Surface Protein 2.";
RL Infect. Immun. 0:0-0(2001).
DR EMBL; AF104716; AAG47624.1; -
DR NON_TER 1
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 11260 MW; C62973C6D0B191D1 CRC64;

Query Match 100.0%; Score 41; DB 5; Length 117;
Best Local Similarity 100.0%; Pred. No. 0.47; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

QY 1 SNTFINNA 8
Db |11111111|
Db 10 SNTFINNA 17

RESULT 6
Q9BJP9 PRELIMINARY; PRT; 117 AA.
AC Q9BJP9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]

OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OKSIBIL SERIES 4, PATIENT 3;
RA MEDLINE=98084480; PubMed=9423864;
RA Eisen D., Billman-Jacobe H., Marshall V.F., Fryauff D., Coppel R.L.;
RT "Temporal variation of the merozoite surface protein-2 gene of
RT Plasmodium falciparum.";
RL Infect. Immun. 66:239-246(1998).
DR EMBL; U72950; AAC02230.1; -
DR InterPro; IPR001136; MSA_2.
DR Pfam; PF00985; MSA_2; 1.
KW Merozoite.
FT NON_TER 1
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 11470 MW; 4501F8D2AF843585 CRC64;

Query Match 100.0%; Score 41; DB 5; Length 111;
Best Local Similarity 100.0%; Pred. No. 0.45; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

QY 1 SNTFINNA 8
Db |11111111|
Db 6 SNTFINNA 13

RESULT 3
Q9BJN8 PRELIMINARY; PRT; 113 AA.
AC Q9BJN8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329575; AAK19394.1; -
DR NON_TER 1
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 11563 MW; C00196E73EAFE7B7 CRC64;

Query Match 100.0%; Score 41; DB 5; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.45; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

QY 1 SNTFINNA 8
Db |11111111|
Db 10 SNTFINNA 17

RESULT 4
Q9GQZ6 PRELIMINARY; PRT; 116 AA.
AC Q9GQZ6;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN (FRAGMENT).
GN MSP2.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN [1]

RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329564; AAK19383.1; -
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 10433 MW; 25D364D909E154B5 CRC64;

Query Match 100.0%; Score 41; DB 5; Length 117;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
Db |11111111
Db 6 SNTFINNA 13

RESULT 7
Q9BJP2
ID Q9BJP2 PRELIMINARY; PRT; 118 AA.
AC Q9BJP2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329571; AAK19390.1; -
FT NON_TER 1
FT NON_TER 118
SQ SEQUENCE 118 AA; 11750 MW; 1D6BED72DE9A0651 CRC64;

Query Match 100.0%; Score 41; DB 5; Length 118;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
Db |11111111
Db 9 SNTFINNA 16

RESULT 8
Q9GQZ0
ID Q9GQZ0 PRELIMINARY; PRT; 119 AA.
AC Q9GQZ0;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN (FRAGMENT).
GN MSP2.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VN15;
RA Weisman S., Wang L., Billman-Jacobe H., Hanh Nhan D., Richie T.L.,
RA Coppel R.L.;
RT "Antibody responses to infections with strains of Plasmodium
RT falciparum expressing diverse forms of Merozoite Surface Protein 2.";

RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329564; AAK19383.1; -
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 10433 MW; 25D364D909E154B5 CRC64;

Query Match 100.0%; Score 41; DB 5; Length 117;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
Db |11111111
Db 6 SNTFINNA 13

RESULT 7
Q9BJP2
ID Q9BJP2 PRELIMINARY; PRT; 118 AA.
AC Q9BJP2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329571; AAK19390.1; -
FT NON_TER 1
FT NON_TER 118
SQ SEQUENCE 118 AA; 11750 MW; 1D6BED72DE9A0651 CRC64;

Query Match 100.0%; Score 41; DB 5; Length 118;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
Db |11111111
Db 9 SNTFINNA 16

RESULT 8
Q9GQZ0
ID Q9GQZ0 PRELIMINARY; PRT; 119 AA.
AC Q9GQZ0;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN (FRAGMENT).
GN MSP2.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VN15;
RA Weisman S., Wang L., Billman-Jacobe H., Hanh Nhan D., Richie T.L.,
RA Coppel R.L.;
RT "Antibody responses to infections with strains of Plasmodium
RT falciparum expressing diverse forms of Merozoite Surface Protein 2.";

RL Infect. Immun. 0:0-0(2001).
DR EMBL; AF104698; AAG47606.1; -
DR InterPro: IPR001136; MSA_2.
DR Pfam: PF00985; MSA_2; 1.
FT NON_TER 1
FT NON_TER 119
SQ SEQUENCE 119 AA; 12070 MW; E379E605000E61E1 CRC64;

Query Match 100.0%; Score 41; DB 5; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
Db |11111111
Db 10 SNTFINNA 17

RESULT 9
Q9BJP7
ID Q9BJP7 PRELIMINARY; PRT; 119 AA.
AC Q9BJP7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329566; AAK19385.1; -
FT NON_TER 1
FT NON_TER 119
SQ SEQUENCE 119 AA; 10408 MW; 6D420B2E261A49BB CRC64;

Query Match 100.0%; Score 41; DB 5; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
Db |11111111
Db 2 SNTFINNA 9

RESULT 10
Q9GOY0
ID Q9GOY0 PRELIMINARY; PRT; 120 AA.
AC Q9GOY0;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE MEROZOITE SURFACE PROTEIN (FRAGMENT).
GN MSP2.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VN25;
RA Weisman S., Wang L., Billman-Jacobe H., Hanh Nhan D., Richie T.L.,
RA Coppel R.L.;
RT "Antibody responses to infections with strains of Plasmodium
RT falciparum expressing diverse forms of Merozoite Surface Protein 2.";

FT NON_TER 120 120
SQ SEQUENCE 120 AA; 11456 MW; D8420C7AC4AC3ED5 CRC64;

Query Match 100.0%; Score 41; DB 5; Length 120;
Best Local Similarity 100.0%; Pred. No. 0.48; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
Db 4 SNTFINNA 11
|||||

RESULT 11
Q9BJN7 PRELIMINARY; PRT; 122 AA.
AC Q9BJN7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE -MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Cloonan N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329576; AAK19395.1; -.
DR NON_TER 1
FT NON_TER 122
SQ SEQUENCE 122 AA; 12497 MW; C3601DB85982D1CB CRC64;

Query Match 100.0%; Score 41; DB 5; Length 122;
Best Local Similarity 100.0%; Pred. No. 0.49; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
Db 6 SNTFINNA 13
|||||

RESULT 12
Q9GQY9 PRELIMINARY; PRT; 125 AA.
AC Q9GQY9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN (FRAGMENT).
GN MSP2.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VN16;
RA Weisman S., Wang L., Billman-Jacobe H., Hanh Nhan D., Richie T.L.,
RA Coppel R.L.;
RT "Antibody responses to infections with strains of Plasmodium
RT falciparum expressing diverse forms of Merozoite Surface Protein 2.";
RL Infect. Immun. 0:0-0(2001).
DR EMBL; AF104699; AAG47607.1; -.
DR InterPro; IPR001136; MSA_2.
DR Pfam; PF00985; MSA_2; 1.
DR NON_TER 1
FT NON_TER 125
SQ SEQUENCE 125 AA; 12605 MW; CFA34C33DEF8ABC6 CRC64;

Query Match 100.0%; Score 41; DB 5; Length 125;
Best Local Similarity 100.0%; Pred. No. 0.5; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
Db 16 SNTFINNA 23
|||||

RESULT 13
Q9GQZ4 PRELIMINARY; PRT; 128 AA.
AC Q9GQZ4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN (FRAGMENT).
GN MSP2.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VN11;
RA Weisman S., Wang L., Billman-Jacobe H., Hanh Nhan D., Richie T.L.,
RA Coppel R.L.;
RT "Antibody responses to infections with strains of Plasmodium
RT falciparum expressing diverse forms of Merozoite Surface Protein 2.";
RL Infect. Immun. 0:0-0(2001).
DR EMBL; AF104694; AAG47602.1; -.
DR InterPro; IPR001136; MSA_2.
DR Pfam; PF00985; MSA_2; 1.
DR NON_TER 1
FT NON_TER 128
SQ SEQUENCE 128 AA; 13038 MW; 4FC1DEFD3402065D CRC64;

Query Match 100.0%; Score 41; DB 5; Length 128;
Best Local Similarity 100.0%; Pred. No. 0.51; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8
Db 10 SNTFINNA 17
|||||

RESULT 14
Q9GQZ5 PRELIMINARY; PRT; 129 AA.
AC Q9GQZ5;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN (FRAGMENT).
GN MSP2.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VN10;
RA Weisman S., Wang L., Billman-Jacobe H., Hanh Nhan D., Richie T.L.,
RA Coppel R.L.;
RT "Antibody responses to infections with strains of Plasmodium
RT falciparum expressing diverse forms of Merozoite Surface Protein 2.";
RL Infect. Immun. 0:0-0(2001).
DR EMBL; AF104693; AAG47601.1; -.
DR InterPro; IPR001136; MSA_2.
DR Pfam; PF00985; MSA_2; 1.
DR NON_TER 1
FT NON_TER 129
SQ SEQUENCE 129 AA; 13080 MW; EBC42289BC7277E5 CRC64;

Query Match 100.0%; Score 41; DB 5; Length 129;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNTFINNA 8
| | | | |
Db 11 SNTFINNA 18

RESULT 15

Q94664
ID Q94664 PRELIMINARY; PRT; 131 AA.
AC Q94664;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).
GN MSP-2.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID-5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OKSIBIL SERIES 2, PATIENT 4;
RX MEDLINE-98084480; PubMed-9423864;
RA Eisen D., Billman-Jacobe H., Marshall V.F., Fryauff D., Coppel R.L.;
RT "Temporal variation of the merozoite surface protein-2 gene of
Plasmodium falciparum.";
RL Infect. Immun. 66:239-246(1998).
DR EMBL; U72948; AAC02228.1; -.
DR InterPro: IPR001136; MSA_2.
DR Pfam: PF00985; MSA_2; 1.
KW Merozoite.
FT NON_TER 1 1
FT NON_TER 131 131
SQ SEQUENCE 131 AA; 13208 MW; F5C707BB681691C5 CRC64;

Query Match 100.0%; Score 41; DB 5; Length 131;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNTFINNA 8
| | | | |
Db 7 SNTFINNA 14

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:21:47 ; Search time 310.82 Seconds
(without alignments)
1.907 Million cell updates/sec

Title: US-09-763-397A-15
Perfect score: 52
Sequence: 1 GQGHMHG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_1101.*
- 1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
 - 2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
 - 3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
 - 4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
 - 5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
 - 6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
 - 7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
 - 8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
 - 9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
 - 10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
 - 11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
 - 12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
 - 13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
 - 14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
 - 15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
 - 16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
 - 17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
 - 18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
 - 19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
 - 20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
 - 21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
 - 22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	8	21 AAY70291	Plasmodium falcipa
2	52	100.0	264	9 AAP80365	Sequence of Ag513
3	52	100.0	272	11 AAR05876	Merozoite surface a
4	52	100.0	272	11 AAR05876	Plasmodium falcipa
5	52	100.0	280	11 AAR05879	Merozoite surface a
6	52	100.0	287	11 AAR05877	Merozoite surface a
7	52	100.0	300	11 AAR05878	Merozoite surface a
8	52	100.0	350	21 AAT70278	Recombinant vacci
9	41	78.8	294	21 AAG29680	Arabidopsis thalia
10	41	78.8	374	21 AAG29679	Arabidopsis thalia
11	41	78.8	408	21 AAG29678	Arabidopsis thalia

12	40	76.9	85	13	AAR26414	Food additive prot
13	40	76.9	256	21	AAV75110	Neisseria meningit
14	40	76.9	257	21	AAV75109	Neisseria meningit
15	39	75.0	102	15	AAR46087	Zinc resistant lik
16	39	75.0	123	22	AAG76052	Human colon cancer
17	39	75.0	376	22	AAW38896	Human polypeptide
18	39	75.0	388	22	AAW40682	Human polypeptide
19	39	75.0	1463	20	AAW99482	Murine NCOA-2 prot
20	38	73.1	137	21	AAG27440	Arabidopsis thalia
21	38	73.1	176	22	AAG63784	Human prostate can
22	38	73.1	217	21	AAG27439	Arabidopsis thalia
23	38	73.1	220	21	AAG27438	Arabidopsis thalia
24	38	73.1	282	21	AAG27318	Arabidopsis thalia
25	38	73.1	336	21	AAG49579	Arabidopsis thalia
26	38	73.1	423	21	AAG27317	Arabidopsis thalia
27	38	73.1	425	21	AAG27316	Arabidopsis thalia
28	38	73.1	477	21	AAG49578	Arabidopsis thalia
29	38	73.1	479	21	AAG49577	Arabidopsis thalia
30	37	71.2	241	21	AAV52698	Drosophila melanog
31	37	71.2	293	21	AAV86316	Human secreted pro
32	37	71.2	307	21	AAV86410	Human gene 27-enco
33	37	71.2	309	21	AAG06085	Arabidopsis thalia
34	37	71.2	309	21	AAG22955	Arabidopsis thalia
35	37	71.2	330	21	AAG22265	Arabidopsis thalia
36	37	71.2	330	21	AAG43480	Arabidopsis thalia
37	37	71.2	344	21	AAG22264	Arabidopsis thalia
38	37	71.2	344	21	AAG43479	Arabidopsis thalia
39	37	71.2	373	22	AAW40116	Human polypeptide
40	37	71.2	375	22	AAW41902	Human polypeptide
41	37	71.2	389	21	AAG06084	Arabidopsis thalia
42	37	71.2	389	21	AAG22954	Arabidopsis thalia
43	37	71.2	398	21	AAG22263	Arabidopsis thalia
44	37	71.2	398	21	AAG43478	Arabidopsis thalia
45	37	71.2	425	21	AAG06063	Arabidopsis thalia

ALIGNMENTS

RESULT 1

AAV70291
ID AAY70291 standard; peptide; 8 AA.

XX AAY70291;

XX 06-JUN-2000 (first entry)

XX Plasmodium falciparum MSP-2 antigenic epitope, P544.

XX Recombinant protein; CDC/NIAIDVAC-1; multivalent; malaria; vaccine;

XX T-cell epitope; tetanus toxoid; antigenic epitope; treatment; SSP-2;

XX circumsporozoite protein; CSP; sporozoite surface protein-1; MSP-1; MSP-2;

XX liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;

XX apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;

XX EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;

XX Pf27; antiparasitic; prevention; anti-CDC/NIAIDVAC-1 antibody.

XX Plasmodium falciparum.

XX WO200011179-A1.

XX 02-MAR-2000.

XX 19-AUG-1999; 99WO-US18869.

XX 21-AUG-1998; 98US-0097703.

XX (NIAID-) NAT INST IMMUNOLOGY.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Lal AA, Shi YP, Hasnain SE;

XX WPI; 2000-237654/20.

XX Novel recombinant protein as vaccine for treating malarial infection
 PT comprises antigenic peptides obtained from different stages of
 PT plasmodium falciparum life cycle
 XX
 PS Claim 2; Page 17; 52pp; English.
 XX
 CC The present sequence is the antigenic epitope P544, derived from
 CC merozoite surface protein-2 (MSP-2) of the asexual blood stage of
 CC Plasmodium falciparum. It is used in the construction of recombinant
 CC protein CDC/NIIMALVAC-1, which is a multivalent, multistage malarial
 CC vaccine. The recombinant protein comprises, melittin signal peptide,
 CC (His)₆ tag, r-cell epitope from tetanus toxoid and 21 antigenic epitopes
 CC from circumsporozoite protein (CSP), sporozoite surface protein-2
 CC (SSP-2), liver stage antigen-1 (LSA-1), merozoite surface protein-1
 CC (MSP-1), MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding
 CC antigen-175 (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete
 CC specific antigen, Pf27. These epitopes were obtained at different stages
 CC of the life cycle of P. falciparum. CDC/NIIMALVAC-1 vaccine has
 CC antiparasitic activity and can be used for treatment and prevention of
 CC malarial infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for
 CC detecting P. falciparum in biological samples.

Query Match 100.0%; Score 52; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 4.3e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHMG 8
 ID 1 gqghmhmg 8

RESULT 2

AAP80365
 ID AAP80365 standard; protein; 264 AA.

XX
 AC AAP80365;

XX 10-DEC-1990 (first entry)

XX Sequence of Ag513 a putative merozoite surface antigen of
 DE Plasmodium falciparum.

XX Malaria therapy; diagnosis.

XX Plasmodium falciparum.

XX Key Location/Qualifiers

FT Peptide 6..20

FT /note="Core of a signal sequence"

FT Region 60..123

FT /note="It contains two identical copies of a 32 AA unit
 arranged in tandem"

FT Region 248..264

FT /note="Hydrophobic sequence presumed to be an integral
 membrane protein anchor sequence"

XX W08800595-A.

XX 28-JAN-1988.

XX 17-JUL-1987; 87WO-AU00227.

XX 24-MAR-1987; 87AU-0001048.

XX (SARA-) SARAWANE PTY LTD (EPPI/).

XX Epping RJ, Ramasamy R, Smythe JA, Anders RF, Coppel RL, Geysen HM;

PI Saul AJ;

XX

DR WPI; 1988-036431/05.
 DR N-PSDB; AAN80340.

XX Merozoite surface antigen of Plasmodium falciparum -
 PT used for active immunisation or for producing monoclonal
 PT antibodies for passive immunisation, purificn. or diagnosis

XX Claim 3; Fig 6; 59pp; English.

XX An antigen having the AA sequence in AAP80365 or an antigenic fragment
 CC thereof is claimed. Also claimed is a recombinant DNA molecule
 CC comprising all or a portion of a nucleotide sequence which is capable of
 CC being expressed as polypeptide having the antigenicity of the above
 CC antigen or an antigen fragment. Monoclonal antibodies produced using the
 CC antigens prevent release of merozoites into the blood stream of an
 CC infected individual. The antigens can also be used for actively
 CC immunising a host against P. falciparum. The antibodies can also be used
 CC for purificn. or detection of P. falciparum.

XX Sequence 264 AA;

Query Match 100.0%; Score 52; DB 9; Length 264;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHMG 8

DB 198 gqghmhmg 205

RESULT 3

AAR05876
 ID AAR05876 standard; protein; 272 AA.

XX
 AC AAR05876;

XX 11-DEC-1991 (first entry)

XX Merozoite surface antigen-2 of the P. falciparum 3D7 isolate.

XX MSA2; Indochina 1 isolate; malaria.

XX Plasmodium falciparum.

XX Key Location/Qualifiers

FT Active-site 27..34

FT /label= Antigenic Sequence

FT /note= "Claim 12"

FT Active-site 207..214

FT /label= Antigenic sequence

FT /note= "Claim 12"

FT Active-site 221..228

FT /label= Antigenic sequence

FT /note= "Claim 12"

FT Region 101..108

FT /label= Thr repeat sequence

XX W09002752-A.

XX 22-MAR-1990.

XX 12-SEP-1989; 89WO-AU00388.

XX 24-AUG-1989; 89AU-0005962.

XX 12-SEP-1988; 88AU-0000382.

XX (SARA-) SARAWANE PTY LTD.

XX Smythe JA, Anders RF, Coppel RL, Saul AJ, Jones GL;

PI Irving DO, Dyer SL;

XX WPI; 1990-115951/15.

DR

SQ Sequence 280 AA;

Query Match 100.0%; Score 52; DB 11; Length 280;

Best Local Similarity 100.0%; Pred. No. 0.16; 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8

Db 214 gqghmhg 221

RESULT 6

AAR05877

ID AAR05877 standard; Protein; 287 AA.

XX AC AAR05877;

XX DT 11-DEC-1991 (first entry)

XX DE Merozoite surface antigen-2 of the P.falciparum Indochina 1 isolate.

XX KW MSA2; 3D7 isolate; malaria.

XX OS Plasmodium falciparum.

XX FH Key Location/Qualifiers

XX FT Region 127..134

XX FT /label= Thr repeat region

XX PN WO9002752-A.

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

ID AAR05878 standard; Protein; 300 AA.

XX AC

XX AAR05878;

XX DT 11-DEC-1991 (first entry)

XX DE Merozoite surface antigen-2 of the P.falciparum MAD71 isolate.

XX KW MSA2; Indochina 1 isolate; malaria.

XX XX Plasmodium falciparum.

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

ID

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

ID

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

ID

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

ID

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158222.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161992.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147204.	PR	29-OCT-1999;	99US-0162142.
PR	04-AUG-1999;	99US-0147302.	Query Match 78.8%; Score 41; DB 21; Length 294;		
PR	05-AUG-1999;	99US-0147192.	Best Local Similarity 100.0%; Pred. No. 12;		
PR	05-AUG-1999;	99US-0147260.	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149368.			
PR	17-AUG-1999;	99US-0149175.			
PR	18-AUG-1999;	99US-0149426.			
PR	20-AUG-1999;	99US-0149722.			
PR	20-AUG-1999;	99US-0149723.			
PR	23-AUG-1999;	99US-0149929.			
PR	23-AUG-1999;	99US-0149902.			
PR	23-AUG-1999;	99US-0149930.			
PR	26-AUG-1999;	99US-0150566.			
PR	26-AUG-1999;	99US-0150884.			
PR	27-AUG-1999;	99US-0151065.			
PR	27-AUG-1999;	99US-0151066.			
PR	27-AUG-1999;	99US-0151080.			
PR	30-AUG-1999;	99US-0151303.			
PR	31-AUG-1999;	99US-0151438.			
			Oy 3 HGHMHG 8		
			Db 123 hghmhg 128		
			RESULT 10		
			AAG29679		
			ID AAG29679 standard; Protein; 374 AA.		
			XX		
			AC AAG29679;		
			XX		
			XX		
			DT 17-OCT-2000 (first entry)		
			XX		
			DE Arabidopsis thaliana protein fragment SEQ ID NO: 35354.		
			XX		
			KW Protein identification; signal transduction pathway; metabolic pathway;		
			KW hybridisation assay; genetic mapping; gene expression control; promoter;		
			KW termination sequence.		
			XX		
			OS Arabidopsis thaliana.		
			XX		
			PN EP1033405-A2.		

XX	06-SEP-2000.		
PD			
XX			
PF	25-FEB-2000; 2000EP-0301439.		
XX			
PR	25-FEB-1999; 99US-0121825.	02-JUL-1999;	99US-0142055.
PR	05-MAR-1999; 99US-0123180.	PR 06-JUL-1999;	99US-0142390.
PR	09-MAR-1999; 99US-0123548.	PR 08-JUL-1999;	99US-0142803.
PR	23-MAR-1999; 99US-0125788.	PR 09-JUL-1999;	99US-0142920.
PR	25-MAR-1999; 99US-0126264.	PR 12-JUL-1999;	99US-0142977.
PR	29-MAR-1999; 99US-0126785.	PR 13-JUL-1999;	99US-0143542.
PR	01-APR-1999; 99US-0127462.	PR 14-JUL-1999;	99US-0143624.
PR	06-APR-1999; 99US-0128234.	PR 15-JUL-1999;	99US-0144005.
PR	08-APR-1999; 99US-0128714.	PR 16-JUL-1999;	99US-0144085.
PR	16-APR-1999; 99US-0129845.	PR 19-JUL-1999;	99US-0144525.
PR	19-APR-1999; 99US-0130077.	PR 19-JUL-1999;	99US-0144331.
PR	21-APR-1999; 99US-0130449.	PR 19-JUL-1999;	99US-0144333.
PR	23-APR-1999; 99US-0130510.	PR 19-JUL-1999;	99US-0144335.
PR	28-APR-1999; 99US-0130891.	PR 20-JUL-1999;	99US-0144352.
PR	30-APR-1999; 99US-0131449.	PR 20-JUL-1999;	99US-0144632.
PR	30-APR-1999; 99US-0132048.	PR 20-JUL-1999;	99US-0144884.
PR	04-MAY-1999; 99US-0132484.	PR 21-JUL-1999;	99US-0144814.
PR	05-MAY-1999; 99US-0132485.	PR 21-JUL-1999;	99US-0145086.
PR	06-MAY-1999; 99US-0132486.	PR 22-JUL-1999;	99US-0145088.
PR	06-MAY-1999; 99US-0132487.	PR 22-JUL-1999;	99US-0145087.
PR	07-MAY-1999; 99US-0132863.	PR 22-JUL-1999;	99US-0145089.
PR	11-MAY-1999; 99US-0134256.	PR 23-JUL-1999;	99US-0145192.
PR	14-MAY-1999; 99US-0134218.	PR 23-JUL-1999;	99US-0145145.
PR	14-MAY-1999; 99US-0134219.	PR 23-JUL-1999;	99US-0145218.
PR	14-MAY-1999; 99US-0134221.	PR 26-JUL-1999;	99US-0145224.
PR	14-MAY-1999; 99US-0134370.	PR 27-JUL-1999;	99US-0145276.
PR	18-MAY-1999; 99US-0134768.	PR 27-JUL-1999;	99US-0145913.
PR	19-MAY-1999; 99US-0134941.	PR 27-JUL-1999;	99US-0145918.
PR	20-MAY-1999; 99US-0135124.	PR 28-JUL-1999;	99US-0145919.
PR	21-MAY-1999; 99US-0135353.	PR 02-AUG-1999;	99US-0145951.
PR	24-MAY-1999; 99US-0135629.	PR 02-AUG-1999;	99US-0146386.
PR	25-MAY-1999; 99US-0136021.	PR 02-AUG-1999;	99US-0146388.
PR	27-MAY-1999; 99US-0136392.	PR 03-AUG-1999;	99US-0147038.
PR	28-MAY-1999; 99US-0136782.	PR 04-AUG-1999;	99US-0147204.
PR	01-JUN-1999; 99US-0137222.	PR 04-AUG-1999;	99US-0147302.
PR	03-JUN-1999; 99US-0137528.	PR 05-AUG-1999;	99US-0147192.
PR	04-JUN-1999; 99US-0137502.	PR 05-AUG-1999;	99US-0147260.
PR	07-JUN-1999; 99US-0137724.	PR 06-AUG-1999;	99US-0147303.
PR	08-JUN-1999; 99US-0138094.	PR 06-AUG-1999;	99US-0147416.
PR	10-JUN-1999; 99US-0138540.	PR 09-AUG-1999;	99US-0147493.
PR	10-JUN-1999; 99US-0138847.	PR 09-AUG-1999;	99US-0147935.
PR	14-JUN-1999; 99US-0139119.	PR 10-AUG-1999;	99US-0148171.
PR	16-JUN-1999; 99US-0139452.	PR 11-AUG-1999;	99US-0148319.
PR	17-JUN-1999; 99US-0139453.	PR 12-AUG-1999;	99US-0148341.
PR	17-JUN-1999; 99US-0139492.	PR 13-AUG-1999;	99US-0148565.
PR	18-JUN-1999; 99US-0139454.	PR 13-AUG-1999;	99US-0148684.
PR	18-JUN-1999; 99US-0139455.	PR 16-AUG-1999;	99US-0149368.
PR	18-JUN-1999; 99US-0139456.	PR 17-AUG-1999;	99US-0149175.
PR	18-JUN-1999; 99US-0139457.	PR 18-AUG-1999;	99US-0149426.
PR	18-JUN-1999; 99US-0139458.	PR 20-AUG-1999;	99US-0149722.
PR	18-JUN-1999; 99US-0139459.	PR 20-AUG-1999;	99US-0149723.
PR	18-JUN-1999; 99US-0139460.	PR 23-AUG-1999;	99US-0149929.
PR	18-JUN-1999; 99US-0139461.	PR 23-AUG-1999;	99US-0149902.
PR	18-JUN-1999; 99US-0139462.	PR 23-AUG-1999;	99US-0149930.
PR	18-JUN-1999; 99US-0139463.	PR 26-AUG-1999;	99US-0150566.
PR	18-JUN-1999; 99US-0139750.	PR 26-AUG-1999;	99US-0150884.
PR	18-JUN-1999; 99US-0139763.	PR 27-AUG-1999;	99US-0151065.
PR	21-JUN-1999; 99US-0139817.	PR 27-AUG-1999;	99US-0151066.
PR	22-JUN-1999; 99US-0139899.	PR 27-AUG-1999;	99US-0151080.
PR	23-JUN-1999; 99US-0140353.	PR 30-AUG-1999;	99US-0151303.
PR	23-JUN-1999; 99US-0140354.	PR 31-AUG-1999;	99US-0151438.
PR	24-JUN-1999; 99US-0140695.	PR 01-SEP-1999;	99US-0151930.
PR	28-JUN-1999; 99US-0140823.	PR 07-SEP-1999;	99US-0152363.
PR	29-JUN-1999; 99US-0140591.	PR 10-SEP-1999;	99US-0153070.
PR	30-JUN-1999; 99US-0141287.	PR 13-SEP-1999;	99US-0153758.
PR	01-JUL-1999; 99US-0141842.	PR 15-SEP-1999;	99US-0154018.
PR	01-JUL-1999; 99US-0142154.	PR 16-SEP-1999;	99US-0154039.
PR		PR 20-SEP-1999;	99US-0154779.

PR	22-SEP-1999;	99US-0155139.	PR	09-MAR-1999;	99US-0123548.
PR	23-SEP-1999;	99US-0155486.	PR	23-MAR-1999;	99US-0125788.
PR	24-SEP-1999;	99US-0155659.	PR	25-MAR-1999;	99US-0126264.
PR	28-SEP-1999;	99US-0156458.	PR	29-MAR-1999;	99US-0126785.
PR	29-SEP-1999;	99US-0156596.	PR	01-APR-1999;	99US-0127462.
PR	04-OCT-1999;	99US-0157117.	PR	06-APR-1999;	99US-0128234.
PR	05-OCT-1999;	99US-0157753.	PR	08-APR-1999;	99US-0128714.
PR	06-OCT-1999;	99US-0157865.	PR	16-APR-1999;	99US-0129845.
PR	07-OCT-1999;	99US-0158029.	PR	19-APR-1999;	99US-0130077.
PR	08-OCT-1999;	99US-0158232.	PR	21-APR-1999;	99US-0130449.
PR	12-OCT-1999;	99US-0158369.	PR	23-APR-1999;	99US-0130510.
PR	13-OCT-1999;	99US-0159293.	PR	23-APR-1999;	99US-0130891.
PR	13-OCT-1999;	99US-0159294.	PR	28-APR-1999;	99US-0131449.
PR	13-OCT-1999;	99US-0159295.	PR	30-APR-1999;	99US-0132048.
PR	14-OCT-1999;	99US-0159329.	PR	30-APR-1999;	99US-0132407.
PR	14-OCT-1999;	99US-0159330.	PR	04-MAY-1999;	99US-0132484.
PR	14-OCT-1999;	99US-0159331.	PR	05-MAY-1999;	99US-0132485.
PR	14-OCT-1999;	99US-0159637.	PR	06-MAY-1999;	99US-0132486.
PR	14-OCT-1999;	99US-0159638.	PR	06-MAY-1999;	99US-0132487.
PR	18-OCT-1999;	99US-0159594.	PR	07-MAY-1999;	99US-0132863.
PR	21-OCT-1999;	99US-0160741.	PR	11-MAY-1999;	99US-0134256.
PR	21-OCT-1999;	99US-0160767.	PR	14-MAY-1999;	99US-0134218.
PR	21-OCT-1999;	99US-0160788.	PR	14-MAY-1999;	99US-0134219.
PR	21-OCT-1999;	99US-0160770.	PR	14-MAY-1999;	99US-0134221.
PR	21-OCT-1999;	99US-0160814.	PR	14-MAY-1999;	99US-0134370.
PR	21-OCT-1999;	99US-0160815.	PR	18-MAY-1999;	99US-0134768.
PR	22-OCT-1999;	99US-0160980.	PR	19-MAY-1999;	99US-0134941.
PR	22-OCT-1999;	99US-0160981.	PR	20-MAY-1999;	99US-0135124.
PR	22-OCT-1999;	99US-0160989.	PR	21-MAY-1999;	99US-0135353.
PR	22-OCT-1999;	99US-0161404.	PR	24-MAY-1999;	99US-0135829.
PR	25-OCT-1999;	99US-0161405.	PR	25-MAY-1999;	99US-0136021.
PR	25-OCT-1999;	99US-0161406.	PR	27-MAY-1999;	99US-0136392.
PR	26-OCT-1999;	99US-0161359.	PR	28-MAY-1999;	99US-0136782.
PR	26-OCT-1999;	99US-0161360.	PR	01-JUN-1999;	99US-0137222.
PR	26-OCT-1999;	99US-0161361.	PR	03-JUN-1999;	99US-0137528.
PR	28-OCT-1999;	99US-0161920.	PR	04-JUN-1999;	99US-0137502.
PR	28-OCT-1999;	99US-0161952.	PR	07-JUN-1999;	99US-0137724.
PR	28-OCT-1999;	99US-0161993.	PR	08-JUN-1999;	99US-0138094.
PR	29-OCT-1999;	99US-0162142.	PR	10-JUN-1999;	99US-0138540.
PR	29-OCT-1999;	99US-0162142.	PR	10-JUN-1999;	99US-0138847.
PR	29-OCT-1999;	99US-0162142.	PR	14-JUN-1999;	99US-0139119.
PR	29-OCT-1999;	99US-0162142.	PR	14-JUN-1999;	99US-0139119.
PR	29-OCT-1999;	99US-0162142.	PR	16-JUN-1999;	99US-0139452.
PR	29-OCT-1999;	99US-0162142.	PR	16-JUN-1999;	99US-0139453.
PR	29-OCT-1999;	99US-0162142.	PR	16-JUN-1999;	99US-0139453.
PR	29-OCT-1999;	99US-0162142.	PR	17-JUN-1999;	99US-0139492.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139454.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139455.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139456.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139457.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139458.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139459.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139460.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139461.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139462.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139463.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139750.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139750.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139763.
PR	29-OCT-1999;	99US-0162142.	PR	21-JUN-1999;	99US-0139817.
PR	29-OCT-1999;	99US-0162142.	PR	21-JUN-1999;	99US-0139899.
PR	29-OCT-1999;	99US-0162142.	PR	22-JUN-1999;	99US-0139899.
PR	29-OCT-1999;	99US-0162142.	PR	23-JUN-1999;	99US-0140353.
PR	29-OCT-1999;	99US-0162142.	PR	23-JUN-1999;	99US-0140354.
PR	29-OCT-1999;	99US-0162142.	PR	24-JUN-1999;	99US-0140695.
PR	29-OCT-1999;	99US-0162142.	PR	24-JUN-1999;	99US-0140695.
PR	29-OCT-1999;	99US-0162142.	PR	28-JUN-1999;	99US-0140823.
PR	29-OCT-1999;	99US-0162142.	PR	28-JUN-1999;	99US-0140823.
PR	29-OCT-1999;	99US-0162142.	PR	29-JUN-1999;	99US-0140991.
PR	29-OCT-1999;	99US-0162142.	PR	30-JUN-1999;	99US-0141287.
PR	29-OCT-1999;	99US-0162142.	PR	30-JUN-1999;	99US-0141287.
PR	29-OCT-1999;	99US-0162142.	PR	01-JUL-1999;	99US-0141842.
PR	29-OCT-1999;	99US-0162142.	PR	01-JUL-1999;	99US-0142154.
PR	29-OCT-1999;	99US-0162142.	PR	02-JUL-1999;	99US-0142055.
PR	29-OCT-1999;	99US-0162142.	PR	06-JUL-1999;	99US-0142390.
PR	29-OCT-1999;	99US-0162142.	PR	06-JUL-1999;	99US-0142390.
PR	29-OCT-1999;	99US-0162142.	PR	08-JUL-1999;	99US-0142803.
PR	29-OCT-1999;	99US-0162142.	PR	09-JUL-1999;	99US-0142920.
PR	29-OCT-1999;	99US-0162142.	PR	12-JUL-1999;	99US-0142977.
PR	29-OCT-1999;	99US-0162142.	PR	13-JUL-1999;	99US-0143542.
PR	29-OCT-1999;	99US-0162142.	PR	14-JUL-1999;	99US-0143624.

Query Match 78.88; Score 41; DB 21; Length 374;
Best Local Similarity 100.08; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HGHMHG 8
|||||
DB 203 hghmhg 208

RESULT 11
AAG29678
ID AAG29678 standard; Protein; 408 AA.
AC AAG29678;
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35353.
DE Arabidopsis thaliana.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
OS EP1033405-A2.
PW 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-012180.

PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.

PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 78.8%; Score 41; DB 21; Length 408;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HGHMHG 8
Db 237 hghmhg 242

RESULT 12
AAR26414
ID AAR26414 standard; Protein; 85 AA.
XX AAR26414;
XX

DT 16-MAR-1993 (first entry)

DE Food additive protein.

XX Antimycotic agent; thermostable; food additive; body fluid;
KW insect larvae; Diptera; Hymenoptera; sarcophaga peregrina;
KW reverse phase HPLC.

OS Diptera; Hymenoptera.

XX Key Location/Qualifiers
FH Key 1..18
FT Peptide /label= Signal_peptide
FT Protein 19..85
FT /label= Mature_protein

PN JP04266900-A.

XX 22-SEP-1992.

XX 21-FEB-1992; 92JP-0349261.

XX 21-FEB-1991; 91JP-0349261.

PA (AMANO) AMANO PHARM KK.

XX WPI: 1992-363111/44.
 DR N-PSDB; AAQ27738.
 XX
 PT Abtimycotic protein useful as food additive - found in insect larvae
 of diptera or hymenoptera
 XX
 PS Disclosure; Page 8; 9pp; Japanese.
 XX
 CC The sequence given is a protein which is useful as an antimycotic
 agent. The protein has a molecular weight of approx. 13000 and is
 CC thermostable for 5-10 minutes when heated to 100 degrees C. The
 CC protein may be incorporated into food additives for humans and
 CC animals. The protein can be obtained by collecting the body fluids
 CC of insect larvae of Diptera or Hymenoptera (eg. sarcophaga peregrina),
 CC centrifuging the solution and subjecting the supernatant to reverse
 CC phase.HPLC.
 XX
 SQ Sequence 85 AA;
 Query Match 76.9%; Score 40; DB 13; Length 85;
 Best Local Similarity 75.0%; Pred. No. 4.8;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GQGHMHG 8
 |||||
 Db 56 gqghyahg 63
 RESULT 13
 AAY75110
 ID AAY75110 standard; Protein; 256 AA.
 XX
 AC AAY75110;
 XX
 DT 21-MAR-2000 (first entry)
 XX
 DE Neisseria meningitidis ORF 568 protein sequence SEQ ID NO:1694.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KW antibacterial; gene therapy.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO9957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Massignani V, Mora M;
 PI Petersen J, Piza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 XX WPI: 2000-062150/05.
 DR N-PSDB; AAZ53872.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 vaccines and diagnostics

XX Claim 2; Page 873; 1453pp; English.
 PS
 XX
 CC AAZ53015 to AAZ54536. AAZ54577 to AAZ54615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitis and N. gonorrhoea polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX
 SQ Sequence 256 AA;
 Query Match 76.9%; Score 40; DB 21; Length 256;
 Best Local Similarity 85.7%; Pred. No. 16;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GQGHMHG 7
 |||||
 Db 236 gqghrhrh 242
 RESULT 14
 AAY75109
 ID AAY75109 standard; Protein; 257 AA.
 XX
 AC AAY75109;
 XX
 DT 21-MAR-2000 (first entry)
 XX
 DE Neisseria meningitidis ORF 568 protein sequence SEQ ID NO:1692.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KW antibacterial; gene therapy.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO9957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Massignani V, Mora M;
 PI Petersen J, Piza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 XX WPI: 2000-062150/05.
 DR N-PSDB; AAZ53871.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 vaccines and diagnostics

PS Claim 2; Page 872; 1453pp; English.

XX AA253015 to AA254536, AA254577 to AA254615, and AA254616 to AA254617 represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides and polypeptides. AA254537 to AA254576 and AA254616 to AA254617 represent PCR primers used in the exemplification of the present invention. The polypeptides, the polynucleotides, antibodies and compositions of the invention can be used as vaccines, as diagnostic reagents, and as immunogenic compositions. The polypeptides can be used in the manufacture of medicaments for treating or preventing infection due to *Neisseria meningitidis* (e.g. meningitis and septicaemia), to detect the presence of *Neisseria meningitidis*, or to raise antibodies. They may also be used to screen for agonists or antagonists, which may themselves have use as antibacterial agents. The polynucleotides of the invention may also be used in gene therapy protocols.

SQ Sequence 257 AA;

Query Match 76.9%; Score 40; DB 21; Length 257;

Best Local Similarity 85.7%; Pred. No. 16;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GQGHMH 7

Db 237 gqghrh 243

RESULT 15

AA2546087

ID AAR46087 standard; Protein; 102 AA.

XX AAR46087;

XX 19-OCT-1994 (first entry)

XX Zinc resistant like protein.

XX Human cDNA; library; enzyme; protein.

XX Homo sapiens.

XX WO9403599-A.

XX 17-FEB-1994.

XX 04-AUG-1993; 93WO-JP01095.

XX 04-AUG-1992; 92JP-0208077.

XX 13-NOV-1992; 92JP-0327619.

XX 26-FEB-1993; 93JP-0061431.

XX (SAGA) SAGAMI CHEM RES CENTRE.

XX Iwahori A, Kato S, Kato T, Kim N, Oh S, Sekine S;

XX WPI; 1994-065688/08.

XX N-PSDB; AAQ57426.

XX cDNA of human origin and proteins coded by it - which may be expressed by in vivo or in vitro translation using sense RNA or antisense DNA corresponding to the cDNA.

XX Claim 1; Page 41-42; 167pp; Japanese.

XX mRNA expressed in human fibrosarcoma cell line HT-1080 was isolated and used to construct a cDNA library using vector pK1. Clone HP0062 encoding Zn resistant-like protein was isolated.

XX Sequence 102 AA;

Query Match 75.0%; Score 39; DB 15; Length 102;

Best Local Similarity 75.0%; Pred. No. 8.6;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GQGHMHG 8

Db 92 gqghshg 99

Search completed: January 29, 2002, 10:21:47

Job time: 423 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:24:07 ; Search time 133.18 Seconds
(without alignments)
1.352 Million cell updates/sec

Title: US-09-763-397A-15
Perfect score: 52
Sequence: 1 GQGHMHG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	71.2	398	US-09-461-474-17	Sequence 17, Appl
2	36	69.2	430	US-08-945-848-8	Sequence 8, Appl
3	35	67.3	211	US-09-046-894-31	Sequence 31, Appl
4	35	67.3	349	US-09-461-474-12	Sequence 12, Appl
5	35	67.3	355	US-08-758-621-4	Sequence 4, Appl
6	35	67.3	355	US-09-107-858-4	Sequence 4, Appl
7	35	67.3	617	US-08-137-614A-26	Sequence 26, Appl
8	35	67.3	637	US-08-072-064-1	Sequence 1, Appl
9	35	67.3	637	US-08-072-064-4	Sequence 4, Appl
10	35	67.3	637	US-08-072-064-6	Sequence 6, Appl
11	35	67.3	637	US-08-072-064-8	Sequence 8, Appl
12	35	67.3	637	US-08-072-064-8	Sequence 8, Appl
13	34.5	66.3	32	US-08-023-9808-27	Sequence 27, Appl
14	34.5	66.3	32	US-08-486-953A-22	Sequence 22, Appl
15	34	65.4	47	US-09-612-126-4	Sequence 4, Appl
16	34	65.4	62	US-09-612-126-7	Sequence 7, Appl
17	34	65.4	83	US-09-612-126-6	Sequence 6, Appl
18	34	65.4	94	US-09-612-126-10	Sequence 10, Appl
19	34	65.4	179	US-09-612-126-11	Sequence 11, Appl
20	34	65.4	186	US-09-612-126-8	Sequence 8, Appl
21	34	65.4	255	US-09-612-126-1	Sequence 1, Appl
22	34	65.4	688	US-09-141-047-8	Sequence 8, Appl
23	33	63.5	339	US-08-758-621-2	Sequence 2, Appl
24	33	63.5	339	US-09-107-858-2	Sequence 2, Appl
25	33	63.5	382	US-08-878-989-20	Sequence 20, Appl
26	33	63.5	382	US-09-272-796-20	Sequence 20, Appl
27	33	63.5	390	US-08-106-981-4	Sequence 4, Appl

28	33	63.5	448	4	US-09-461-474-8	Sequence 8, Appl
29	33	63.5	474	4	US-09-461-474-10	Sequence 10, Appl
30	32	61.5	17	1	US-08-155-171B-37	Sequence 37, Appl
31	32	61.5	17	2	US-08-435-998-37	Sequence 37, Appl
32	32	61.5	100	1	US-08-202-389-29	Sequence 29, Appl
33	32	61.5	249	1	US-08-155-171B-32	Sequence 32, Appl
34	32	61.5	249	2	US-08-435-998-32	Sequence 32, Appl
35	32	61.5	410	1	US-08-123-343A-5	Sequence 5, Appl
36	32	61.5	410	1	US-08-123-343A-7	Sequence 7, Appl
37	32	61.5	410	4	US-09-431-573-4	Sequence 4, Appl
38	32	61.5	410	4	US-09-431-573-5	Sequence 5, Appl
39	32	61.5	513	1	US-08-202-389-2	Sequence 2, Appl
40	32	61.5	528	2	US-08-466-589-2	Sequence 2, Appl
41	32	61.5	528	3	US-08-700-636-2	Sequence 2, Appl
42	32	61.5	528	3	US-08-467-574-2	Sequence 2, Appl
43	32	61.5	528	4	US-09-217-345-2	Sequence 2, Appl
44	32	61.5	529	1	US-08-496-855A-2	Sequence 2, Appl
45	32	61.5	595	1	US-08-202-389-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-09-461-474-17
; Sequence 17, Application US/09461474
; Patent No. 6278042
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafalski, Antoni
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Plant Metal Transporters
; FILE REFERENCE: BB1303 US NA
; CURRENT APPLICATION NUMBER: US/09/461,474
; EARLIER FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: 60/112,562
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-461-474-17

Query Match 71.2%; Score 37; DB 4; Length 398;
Best Local Similarity 71.4%; Pred. No. 53;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 OHGHMHG 8
Db 217 EHGSHG 223

RESULT 2
US-08-945-848-8
; Sequence 8, Application US/08945848
; Patent No. 5968772
; GENERAL INFORMATION:
; APPLICANT: MATSUSHIRO, Aizo
; TITLE OF INVENTION: PEARL PROTEIN(NACREIN) AND PROCESS FOR
; TITLE OF INVENTION: THE SAME
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,848
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Cavley, Jr., Thomas A.
REGISTRATION NUMBER: 40,944
REFERENCE/DOCKET NUMBER: 19036/34324
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-945-848-8

Query Match 69.2%; Score 36; DB 2; Length 430;
Best Local Similarity 62.5%; Pred. No. 83;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GQHGMHMG 8
|:|:| |
Db 303 GENGKHG 310

RESULT 3
US-09-046-894-31
; Sequence 31, Application US/09046894
; Patent No. 6190857
; GENERAL INFORMATION:
; APPLICANT: Ralph, David
; APPLICANT: An, Gang
; APPLICANT: O'Hara, Mark S.
; APPLICANT: Veltri, Robert
; TITLE OF INVENTION: DIAGNOSIS OF DISEASE STATE USING mRNA
; TITLE OF INVENTION: PROFILES IN PERIPHERAL LEUKOCYTES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/046,894
; FILING DATE: Concurrently Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,576
; FILING DATE: 24-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Nakashima, Richard A.
; REGISTRATION NUMBER: P-42,023
; REFERENCE/DOCKET NUMBER: UROC:014
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:

LENGTH: 211 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-09-046-894-31

Query Match 67.3%; Score 35; DB 4; Length 211;
Best Local Similarity 83.3%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 HGHMHG 8
| | | |
Db 3 HGHSHG 8

RESULT 4
US-09-461-474-12
; Sequence 12, Application US/09461474
; Patent No. 6278042
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafalski, Antoni
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Plant Metal Transporters
; FILE REFERENCE: BB1303 US NA
; CURRENT APPLICATION NUMBER: US/09/461,474
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: 60/112,562
; EARLIER FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Glycine max
US-09-461-474-12

Query Match 67.3%; Score 35; DB 4; Length 349;
Best Local Similarity 83.3%; Pred. No. 97;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 HGHMHG 8
| | | |
Db 136 HGHSHG 141

RESULT 5
US-08-758-621-4
; Sequence 4, Application US/08758621
; Patent No. 5846821
; GENERAL INFORMATION:
; APPLICANT: Gueriot, Mary Lou, and Elide, David J.
; TITLE OF INVENTION: Metal-Regulated Transporters and Uses Therefor
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/758,621
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/018,578

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,848
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Cavley, Jr., Thomas A.
REGISTRATION NUMBER: 40,944
REFERENCE/DOCKET NUMBER: 19036/34324
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-945-848-8

Query Match 69.2%; Score 36; DB 2; Length 430;
Best Local Similarity 62.5%; Pred. No. 83;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GQHGMHMG 8
|:|:| |
Db 303 GENGKHG 310

RESULT 3
US-09-046-894-31
; Sequence 31, Application US/09046894
; Patent No. 6190857
; GENERAL INFORMATION:
; APPLICANT: Ralph, David
; APPLICANT: An, Gang
; APPLICANT: O'Hara, Mark S.
; APPLICANT: Veltri, Robert
; TITLE OF INVENTION: DIAGNOSIS OF DISEASE STATE USING mRNA
; TITLE OF INVENTION: PROFILES IN PERIPHERAL LEUKOCYTES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/046,894
; FILING DATE: Concurrently Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,576
; FILING DATE: 24-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Nakashima, Richard A.
; REGISTRATION NUMBER: P-42,023
; REFERENCE/DOCKET NUMBER: UROC:014
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:

```

; FILING DATE: 29-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Silveri, Jean M.
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: DCI-099CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-758-621-4

```

```

Query Match 67.3%; Score 35; DB 2; Length 355;
Best Local Similarity 83.3%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 3 HGHMHG 8
Db 185 HGHTHG 190

```

```

RESULT 6
US-09-107-858-4
; Sequence 4, Application US/09107858
; Patent No. 6162900
; GENERAL INFORMATION:
; APPLICANT: Guerinot, Mary Lou et al.
; TITLE OF INVENTION: METAL-REGULATED TRANSPORTERS AND USES THEREFOR
; FILE REFERENCE: DCI-099CPDV
; CURRENT APPLICATION NUMBER: US/09/107,858
; CURRENT FILING DATE: 1998-06-30
; EARLIER APPLICATION NUMBER: 08/758,621
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-107-858-4

```

```

Query Match 67.3%; Score 35; DB 4; Length 355;
Best Local Similarity 83.3%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 3 HGHMHG 8
Db 185 HGHTHG 190

```

```

RESULT 7
US-08-137-614A-26
; Sequence 26, Application US/08137614A
; Patent No. 5487976
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Knipple, Douglas C.
; APPLICANT: Henderson, Joseph E.
; TITLE OF INVENTION: Gene Encoding An Insect
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA

```

```

; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,614A
; FILING DATE: 15-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716)263-1636
; TELEFAX: (716)263-1600
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 617 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-137-614A-26

```

```

Query Match 67.3%; Score 35; DB 1; Length 617;
Best Local Similarity 71.4%; Pred. No. 17e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 QHGHMHG 8
Db 436 BHGHGHG 442

```

```

RESULT 8
US-08-072-064-1
; Sequence 1, Application US/08072064
; Patent No. 6008046
; GENERAL INFORMATION:
; APPLICANT: FFRENCH-CONSTANT, RICHARD H.
; APPLICANT: JACKSON, MEYER B.
; TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETER G. CARROLL
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,064
; FILING DATE: 19930602
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 770,881
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OPHD-00574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 637 amino acids

```

TYPE: AMINO ACID
TOPOLOGY: unknown
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Drosophila melanogaster
POSITION IN GENOME:
CHROMOSOME/SEGMENT: III; polytene subregion 66F
MAP POSITION: approximately map unit 26
US-08-072-064-1

Query Match 67.3%; Score 35; DB 3; Length 637;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 QHGHMHG 8
:|||||
Db 457 EHGHGHG 463

RESULT 9
US-08-072-064-4
; Sequence 4, Application US/08072064
; Patent No. 6008046
; GENERAL INFORMATION:
; APPLICANT: FFRENCH-CONSTANT, RICHARD H.
; APPLICANT: JACKSON, MEYER B.
; TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETER G. CARROLL
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 19930602
; APPLICATION NUMBER: US/08/072.064
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OPHD-00574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 637 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-072-064-4

Query Match 67.3%; Score 35; DB 3; Length 637;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 QHGHMHG 8
:|||||
Db 457 EHGHGHG 463

RESULT 10
US-08-072-064-6
; Sequence 6, Application US/08072064
; Patent No. 6008046
; GENERAL INFORMATION:
; APPLICANT: FFRENCH-CONSTANT, RICHARD H.
; APPLICANT: JACKSON, MEYER B.
; TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETER G. CARROLL
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 19930602
; APPLICATION NUMBER: US/08/072.064
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OPHD-00574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 637 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-072-064-6

Query Match 67.3%; Score 35; DB 3; Length 637;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 QHGHMHG 8
:|||||
Db 457 EHGHGHG 463

RESULT 11
US-08-072-064-8
; Sequence 8, Application US/08072064
; Patent No. 6008046
; GENERAL INFORMATION:
; APPLICANT: FFRENCH-CONSTANT, RICHARD H.
; APPLICANT: JACKSON, MEYER B.
; TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETER G. CARROLL
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,064
; FILING DATE: 19930602
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 770,881
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OPHD-00574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 637 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-072-064-8

```

Query Match 67.3%; Score 35; DB 3; Length 637;
 Best Local Similarity 71.4%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 QHGHMHG 8
 Db 457 EHGHHG 463

RESULT 12

```

PCT-US92-08558-1
; Sequence 1, Application PC/TUS9208558
; GENERAL INFORMATION:
; APPLICANT: Cornell Research Foundation, Inc.
; TITLE OF INVENTION: MOLECULAR CLONING AND TRANSFORMATION OF CYCLODIENE RESISTANCE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08558
; FILING DATE: 19921002
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/770,881
; FILING DATE: October 4th 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; REFERENCE/DOCKET NUMBER: CRF D-1052
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 637 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:

```

```

; ORGANISM: Drosophila melanogaster
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: III
; MAP POSITION: approximately map unit 26
PCT-US92-08558-1

```

Query Match 67.3%; Score 35; DB 5; Length 637;
 Best Local Similarity 71.4%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 QHGHMHG 8
 Db 457 EHGHHG 463

RESULT 13

```

US-08-023-980B-27
; Sequence 27, Application US/08023980B
; Patent No. 5843641
; GENERAL INFORMATION:
; APPLICANT: Brown, Robert
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Rosen, Daniel R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS,
; TREATMENT AND PREVENTION OF DISEASES OF CELL DEATH
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 585 Commercial Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-1024
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/023,980B
; FILING DATE: 26-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/177001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/723-4123
; TELEFAX: 617/723-8962
; TELEX:
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-023-980B-27

```

Query Match 66.3%; Score 34.5; DB 2; Length 32;
 Best Local Similarity 66.7%; Pred. No. 10;
 Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 GGHG-HMHG 8
 Db 18 GGHGHHG 26

RESULT 14

```

US-08-486-953A-22
; Sequence 22, Application US/08486953A

```

Patent No. 5849290
; GENERAL INFORMATION:
; APPLICANT: Brown, Robert
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Rosen, Daniel R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS,
; TITLE OF INVENTION: TREATMENT AND PREVENTION OF DISEASES OF CELL DEATH
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FastSeq
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,953A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/204,052
; FILING DATE: 28-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/223002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/428-0200
; TELEFAX: 617/428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-486-953A-22

Query Match 66.3%; Score 34.5; DB 2; Length 32;
Best Local Similarity 66.7%; Pred. No. 10;
Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;
Oy 1 GQHG-HMHG 8
Db 18 GEHGFVHG 26
RESULT 15
US-09-612-126-4
; Sequence 4, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
; FILE REFERENCE: 6056-258 Ctl
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 47

TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human high
; OTHER INFORMATION: molecular weight kininogen light chain amino acids
; OTHER INFORMATION: Gly(456) through Lys(502)
US-09-612-126-4
Query Match 65.4%; Score 34; DB 4; Length 47;
Best Local Similarity 71.4%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy 2 QHGHMHG 8
Db 25 KHGHGHG 31
Search completed: January 29, 2002, 10:24:07
Job time: 513 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:26:42 : Search time 144.96 Seconds
(without alignments)
4.204 Million cell updates/sec

Title: US-09-763-397A-15
Perfect score: 52
Sequence: 1 GOGHMHG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	148	2 B45637	merozoite surface
2	52	100.0	172	2 A45637	merozoite surface
3	52	100.0	180	2 B45613	surface antigen FU
4	52	100.0	256	2 S39311	merozoite surface
5	52	100.0	264	2 A31818	merozoite 45K surf
6	52	100.0	264	2 S53367	merozoite surface
7	52	100.0	272	2 G71618	merozoite surface
8	52	100.0	274	2 A45632	merozoite surface
9	52	100.0	278	2 S39310	merozoite surface
10	52	100.0	280	2 C39112	merozoite 45K surf
11	52	100.0	286	2 B45632	merozoite surface
12	52	100.0	287	2 B39615	merozoite 45K surf
13	52	100.0	300	2 A39112	merozoite 45K surf
14	52	100.0	302	2 A39615	merozoite 45K surf
15	52	100.0	347	2 B39112	merozoite 45K surf
16	41	78.8	374	2 F86243	Z1P4, probable zin
17	41	78.8	549	2 T15506	hypothetical prote
18	40	76.9	85	2 A45969	hemolymph antifung
19	40	76.9	448	1 A56018	transcription fact
20	40	76.9	449	1 S30205	transcription fact
21	40	76.9	451	1 A40168	transcription fact
22	39	75.0	141	2 E86413	hypothetical prote
23	39	75.0	159	2 B82781	peptidyl-prolyl ci
24	39	75.0	160	2 T07180	hypothetical prote
25	39	75.0	241	2 T26522	hypothetical prote
26	39	75.0	391	2 T26756	hypothetical prote
27	39	75.0	410	2 T26757	hypothetical prote
28	39	75.0	411	2 T29235	hypothetical prote
29	39	75.0	436	2 I49714	MHC H-2K/t-w5-link

30	39	75.0	481	2 S60260	stomatin-like prot
31	39	75.0	1127	2 T32404	hypothetical prote
32	38	73.1	171	1 D69986	conserved hypothet
33	38	73.1	203	2 T36240	hypothetical prote
34	38	73.1	431	2 S49821	PRL2 protein - Ara
35	38	73.1	1994	2 D86452	protein F6N18.13 l
36	37	71.2	66	2 B83515	hypothetical prote
37	37	71.2	126	2 E86669	hypothetical prote
38	37	71.2	160	2 A81075	FKBP-type peptidyl
39	37	71.2	161	2 E83541	peptidyl-prolyl ci
40	37	71.2	190	2 D64087	probable peptidylp
41	37	71.2	191	2 C41892	19K membrane-bound
42	37	71.2	196	2 D85999	hypothetical prote
43	37	71.2	196	2 A49987	probable fkbp-type
44	37	71.2	286	2 S07193	chorion protein s3
45	37	71.2	325	2 A75437	cation efflux syst

ALIGNMENTS

RESULT 1

B45637
merozoite surface antigen 2 - malaria parasite (Plasmodium falciparum) (fragments)
C:Species: Plasmodium falciparum
C:Date: 22-Apr-1993 #sequence_revision 02-Jul-1996 #text_change 17-Nov-2000
C:Accession: B45637
R:Snewin, V.A.; Herrera, M.; Sanchez, G.; Scherf, A.; Langsley, G.; Herrera, S.
Mol. Biochem. Parasitol. 49, 265-275, 1991
A:Title: Polymorphism of the alleles of the merozoite surface antigens MSA1 and MSA2
A:Reference number: A45637; MUID:92131063
A:Accession: B45637
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148 <SNE>
A:Experimental source: isolate COL8
A>Note: sequence extracted from NCBI backbone (NCBIN:78806, NCBIP:78808)
C:Superfamily: Epstein-Barr virus nuclear antigen
C:Keywords: surface antigen

Query Match 100.0%; Score 52; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GOGHMHG 8
| | | | | | | |
DB 137 GOGHMHG 144

RESULT 2

A45637
merozoite surface antigen 2 - malaria parasite (Plasmodium falciparum) (fragments)
C:Species: Plasmodium falciparum
C:Date: 22-Apr-1993 #sequence_revision 02-Jul-1996 #text_change 17-Nov-2000
C:Accession: A45637
R:Snewin, V.A.; Herrera, M.; Sanchez, G.; Scherf, A.; Langsley, G.; Herrera, S.
Mol. Biochem. Parasitol. 49, 265-275, 1991
A:Title: Polymorphism of the alleles of the merozoite surface antigens MSA1 and MSA2
A:Reference number: A45637; MUID:92131063
A:Accession: A45637
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-172 <SNE>
A:Experimental source: isolate COL5
A>Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:78805, NCBIP:78809)
C:Superfamily: Epstein-Barr virus nuclear antigen
C:Keywords: surface antigen

Query Match 100.0%; Score 52; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.054;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8
| | | | |
Db 161 GQGHMHG 168

RESULT 3

B45613
surface antigen FUP/SP - malaria parasite (Plasmodium falciparum) (fragments)
C:Species: Plasmodium falciparum
C:Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 17-Nov-2000
C:Accession: B45613
R:Fandeur, T.; Bonnefoy, S.; Mercereau-Puijalat, O.
Mol. Biochem. Parasitol. 47, 167-178, 1991
A:Title: In vivo and in vitro derived Palo Alto lines of Plasmodium falciparum are genotyped
A:Reference number: A45613; MUID:92049549
A:Accession: B45613
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-180 <FAN>
A:Experimental source: merozoite, Uganda Palo Alto strain
A:Note: sequence extracted from NCBI backbone (NCBIP:65036)
C:Superfamily: Epstein-Barr virus nuclear antigen
C:Keywords: surface antigen

Query Match 100.0%; Score 52; DB 2; Length 180;

Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8
| | | | |
Db 169 GQGHMHG 176

RESULT 4

S39311
merozoite surface antigen - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 17-Nov-2000
C:Accession: S39311
R:Ramasamy, R.; Ranasinghe, C.
submitted to the EMBL Data Library, November 1993
A:Description: Cycle ds DNA sequencing of a malaria parasite protein from infected blood
A:Reference number: S39310
A:Accession: S39311
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-256 <RAM>
C:Cross-references: EMBL:X76298; NID:q434997; PID:g836640
C:Superfamily: Epstein-Barr virus nuclear antigen

Query Match 100.0%; Score 52; DB 2; Length 256;

Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8
| | | | |
Db 190 GQGHMHG 197

RESULT 5

A31818
merozoite 45K surface antigen precursor - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 17-Nov-2000
C:Accession: A31818
R:Smythe, J.A.; Coppel, R.L.; Brown, G.V.; Ramasamy, R.; Kemp, D.J.; Anders, R.F.
Proc. Natl. Acad. Sci. U.S.A. 85, 5195-5199, 1988
A:Title: Identification of two integral membrane proteins of Plasmodium falciparum.
A:Reference number: A31818; MUID:88276924

A:Accession: A31818
A:Molecule type: mRNA
A:Residues: 1-264 <SMY>
A:Cross-references: GB:J03828; NID:g160352; PID:g160353
C:Superfamily: Epstein-Barr virus nuclear antigen
C:Keywords: surface antigen
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-264/Product: merozoite 45K surface antigen #status predicted <MAT>

Query Match 100.0%; Score 52; DB 2; Length 264;

Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8
| | | | |
Db 198 GQGHMHG 205

RESULT 6

S55367
merozoite surface antigen 2 - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 17-Nov-2000
C:Accession: S55367
R:Chauhan, V.S.
submitted to the EMBL Data Library, May 1995

A:Reference number: S55367
A:Accession: S55367
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <CHA>
A:Cross-references: EMBL:X87249; NID:g854367; PID:g854368
C:Superfamily: Epstein-Barr virus nuclear antigen
C:Keywords: surface antigen

Query Match 100.0%; Score 52; DB 2; Length 264;

Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8
| | | | |
Db 198 GQGHMHG 205

RESULT 7

G71618
merozoite surface antigen MSP-2 PFB0300c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 17-Nov-2000
C:Accession: G71618; A44950
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.; Perera, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743
A:Accession: G71618
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-272 <GAR>
A:Cross-references: GB:AE001385; GB:AE001362; NID:g3845143; PIDN:AACT1849.1; PID:g384
A:Experimental source: clone 3D7
R:Smythe, J.A.; Peterson, M.G.; Coppel, R.L.; Saul, A.J.; Kemp, D.J.; Anders, R.F.
Mol. Biochem. Parasitol. 39, 227-234, 1990
A:Title: Structural diversity in the 45-kilodalton merozoite surface antigen of Plasmodium falciparum.
A:Reference number: A44950; MUID:90205972
A:Accession: A44950
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-54, T, 56-272 <SMY>
A:Cross-references: GB:M28891; NID:g160458; PID:g160459
C:Genetics:

A:Gene: PFB0300c
C:Superfamily: Epstein-Barr virus nuclear antigen
C:Keywords: surface antigen

Query Match 100.0%; Score 52; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.087;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8

Db 206 GQGHMHG 213

RESULT 8

merozoite surface antigen 2 - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum
C:Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 17-Nov-2000
C:Accession: A45632

R:Marshall, V.M.; Coppel, R.L.; Anders, R.F.; Kemp, D.J.

Mol. Biochem. Parasitol. 50, 181-184, 1992
A:Title: Two novel alleles within subfamilies of the merozoite surface antigen 2 (MSA-2)
A:Reference number: A45632; MUID:92178286

A:Contents: KF1916

A:Accession: A45632

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-274 <MAR>

A:Cross-references: GB:M73810; NID:g160484; PID:g160485

A>Note: sequence extracted from NCBI backbone (NCBIN:85252, NCBIP:85257)

C:Superfamily: Epstein-Barr virus nuclear antigen

C:Keywords: surface antigen

Query Match 100.0%; Score 52; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 0.088;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8

Db 208 GQGHMHG 215

RESULT 9

merozoite surface antigen - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Nov-2000
C:Accession: S39310

R:Ramassany, R.; Ranasinghe, C.

submitted to the EMBL Data Library, November 1993

A:Description: Cycle ds DNA sequencing of a malaria parasite protein from infected blood

A:Reference number: S39310

A:Accession: S39310

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-278 <RAM>

A:Cross-references: EMBL:X76087; NID:g434996; PID:g836639

C:Superfamily: Epstein-Barr virus nuclear antigen

C:Keywords: surface antigen

Query Match 100.0%; Score 52; DB 2; Length 278;
Best Local Similarity 100.0%; Pred. No. 0.089;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8

Db 212 GQGHMHG 219

RESULT 10

C39112

merozoite 45K surface antigen precursor - malaria parasite (Plasmodium falciparum) (1
C:Species: Plasmodium falciparum

C:Date: 30-Aug-1991 #sequence_revision 30-Aug-1991 #text_change 17-Nov-2000

C:Accession: C39112

R:Smythe, J.A.; Coppel, R.L.; Day, K.P.; Martin, R.K.; Oduola, A.M.J.; Kemp, D.J.; An
Proc. Natl. Acad. Sci. U.S.A. 88, 1751-1755, 1991

A:Title: Structural diversity in the Plasmodium falciparum merozoite surface antigen

A:Reference number: A39112; MUID:91156685

A:Accession: C39112

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-280 <SMY>

A:Cross-references: GB:M59768

C:Superfamily: Epstein-Barr virus nuclear antigen

C:Keywords: surface antigen

Query Match 100.0%; Score 52; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.09;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8

Db 214 GQGHMHG 221

RESULT 11

B45632

merozoite surface antigen 2 - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 17-Nov-2000

C:Accession: B45632

R:Marshall, V.M.; Coppel, R.L.; Anders, R.F.; Kemp, D.J.

Mol. Biochem. Parasitol. 50, 181-184, 1992

A:Title: Two novel alleles within subfamilies of the merozoite surface antigen 2 (MSA-2)
A:Reference number: A45632; MUID:92178286

A:Accession: B45632

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-286 <MAR>

A:Experimental source: isolate 311

A>Note: sequence extracted from NCBI backbone (NCBIN:85255, NCBIP:85259)

C:Superfamily: Epstein-Barr virus nuclear antigen

C:Keywords: surface antigen

Query Match 100.0%; Score 52; DB 2; Length 286;
Best Local Similarity 100.0%; Pred. No. 0.092;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8

Db 220 GQGHMHG 227

RESULT 12

B39615

merozoite 45K surface antigen precursor - malaria parasite (Plasmodium falciparum)

N:Alternate names: membrane antigen pf7

C:Species: Plasmodium falciparum

C:Date: 13-Sep-1991 #sequence_revision 13-Sep-1991 #text_change 01-Dec-2000

C:Accession: B39615; A36018; B44950; A45613

R:Fonton, B.; Clark, J.T.; Khan, C.M.A.; Robinson, J.V.; Walliker, D.; Ridley, R.; Sc

Mol. Cell. Biol. 11, 963-971, 1991

A:Title: Structural and antigenic polymorphism of the 35- to 48-kilodalton merozoite

A:Reference number: A39615; MUID:91117264

A:Accession: B39615

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-287 <FEN>

A:Cross-references: EMBL:X53833

A>Note: clone T9-94

Mon Feb 4 15:23:38 2002

C>Date: 13-Sep-1991 #sequence_revision 13-Sep-1991 #text_change 17-Nov-2000
 C:Accession: A39615; S13802
 R:Fenton, B.; Clark, J.T.; Khan, C.M.A.; Robinson, J.V.; Walliker, D.; Ridley, R.; Sc
 Mol. Cell. Biol. 11, 963-971, 1991
 A:Title: Structural and antigenic polymorphism of the 35- to 48-kilodalton merozoite
 A:Reference number: A39615; MUID:91117264
 A:Accession: A39615
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-302 <FEN>
 A:Cross-references: EMBL:X53832
 C:Superfamily: Epstein-Barr virus nuclear antigen
 C:Keywords: surface antigen

R:Elliot, J.F.; Albrecht, G.R.; Gilladoga, A.; Handunnett, S.M.; Neequaye, J.; Lalling
 Proc. Natl. Acad. Sci. U.S.A. 87, 6363-6367, 1990
 A:Title: Genes for Plasmodium falciparum surface antigens cloned by expression in COS ce
 A:Reference number: A36018; MUID:90349616
 A:Accession: A36018
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-287 <ELL>
 A:Cross-references: GB:M28890; NID:g160406; PID:g160407
 R:Smythe, J.A.; Peterson, M.G.; Coppel, R.L.; Saul, A.J.; Kemp, D.J.; Anders, R.F.
 Mol. Biochem. Parasitol. 39, 227-234, 1990
 A:Title: Structural diversity in the 45-kilodalton merozoite surface antigen of Plasmodi
 A:Reference number: A44950; MUID:90205972
 A:Accession: B44950
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-186, 'R', 188-287 <SMY>
 A:Cross-references: GB:M28892; NID:g160488; PID:g160489
 R:Fandeur, T.; Bonnefoy, S.; Mercereau-Puijalon, O.
 Mol. Biochem. Parasitol. 47, 167-178, 1991
 A:Title: In vivo and in vitro derived Palo Alto lines of Plasmodium falciparum are genet
 A:Reference number: A45613; MUID:92049549
 A:Accession: A45613
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 28, 'E', 30-186, 'R', 188-230, 'IH' <FAN>
 A:Experimental source: Uganda Palo Alto strain, merozoite
 A:Note: sequence extracted from NCBI backbone (NCBIP:65035)
 C:Superfamily: Epstein-Barr virus nuclear antigen
 C:Keywords: membrane protein; surface antigen

Query Match 100.0%; Score 52; DB 2; Length 302;
 Best Local Similarity 100.0%; Pred. No. 0.097;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GQGHMHG 8
 |||||||
 Db 236 GQGHMHG 243

RESULT 15
 B39112
 merozoite 45K surface antigen precursor - malaria parasite (Plasmodium falciparum) (1
 C:Species: Plasmodium falciparum
 C>Date: 30-Aug-1991 #sequence_revision 30-Aug-1991 #text_change 17-Nov-2000
 C:Accession: B39112
 R:Smythe, J.A.; Coppel, R.L.; Day, K.P.; Martin, R.K.; Oduola, A.M.J.; Kemp, D.J.; An
 Proc. Natl. Acad. Sci. U.S.A. 88, 1751-1755, 1991
 A:Title: Structural diversity in the Plasmodium falciparum merozoite surface antigen
 A:Reference number: A39112; MUID:91156685
 A:Accession: B39112
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-347 <SMY>
 A:Cross-references: GB:M59767
 C:Superfamily: Epstein-Barr virus nuclear antigen
 C:Keywords: surface antigen

Query Match 100.0%; Score 52; DB 2; Length 287;
 Best Local Similarity 100.0%; Pred. No. 0.092;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GQGHMHG 8
 |||||||
 Db 221 GQGHMHG 228

RESULT 13
 A39112
 merozoite 45K surface antigen precursor - malaria parasite (Plasmodium falciparum) (iso
 C:Species: Plasmodium falciparum
 C>Date: 30-Aug-1991 #sequence_revision 30-Aug-1991 #text_change 17-Nov-2000
 C:Accession: A39112
 R:Smythe, J.A.; Coppel, R.L.; Day, K.P.; Martin, R.K.; Oduola, A.M.J.; Kemp, D.J.; Ander
 Proc. Natl. Acad. Sci. U.S.A. 88, 1751-1755, 1991
 A:Title: Structural diversity in the Plasmodium falciparum merozoite surface antigen 2.
 A:Reference number: A39112; MUID:91156685
 A:Accession: A39112
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-300 <SMY>
 A:Cross-references: GB:M59765
 C:Superfamily: Epstein-Barr virus nuclear antigen
 C:Keywords: surface antigen

Query Match 100.0%; Score 52; DB 2; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.097;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GQGHMHG 8
 |||||||
 Db 234 GQGHMHG 241

RESULT 14
 A39615
 merozoite 45K surface antigen precursor (clone T9-96) - malaria parasite (Plasmodium fal
 C:Species: Plasmodium falciparum

Search completed: January 29, 2002, 10:26:43
 Job time: 654 sec

THIS PAGE BLANK (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:13:44 ; Search time 80.65 Seconds
(without alignments)
3.637 Million cell updates/sec

Title: US-09-763-397A-15
Perfect score: 52
Sequence: 1 GQGHMHG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	52	100.0	262	MSA2_PLAFC	Q99317 plasmodium
2	52	100.0	264	MSA2_PLAFC	P19599 plasmodium
3	52	100.0	264	MSA2_PLAFC	P50498 plasmodium
4	52	100.0	272	MSA2_PLAFC	P50498 plasmodium
5	52	100.0	274	MSA2_PLAFC	P50497 plasmodium
6	52	100.0	276	MSA2_PLAFC	Q99320 plasmodium
7	52	100.0	280	MSA2_PLAFC	Q03643 plasmodium
8	52	100.0	281	MSA2_PLAFC	Q99319 plasmodium
9	52	100.0	286	MSA2_PLAFC	P50496 plasmodium
10	52	100.0	287	MSA2_PLAFC	P19260 plasmodium
11	52	100.0	300	MSA2_PLAFC	Q03644 plasmodium
12	52	100.0	302	MSA2_PLAFC	Q03645 plasmodium
13	52	100.0	302	MSA2_PLAFC	Q03994 plasmodium
14	52	100.0	347	MSA2_PLAFC	Q03646 plasmodium
15	41	78.8	449	CSUP_DROME	Q9v3a4 drosophila
16	40	76.9	85	ANTF_SARPE	Q08617 sarcophaga
17	40	76.9	448	OCT6_HUMAN	Q03052 homo sapien
18	40	76.9	449	OCT6_MOUSE	P21952 mus musculu
19	40	76.9	451	OCT6_RAT	P20267 rattus norv
20	39	75.0	155	KE4_PIG	Q029175 sus scrofa
21	39	75.0	212	SLYD_AERHY	Q07046 aeromonas h
22	39	75.0	469	KE4_HUMAN	Q92504 homo sapien
23	39	75.0	476	KE4_MOUSE	Q31125 mus musculu
24	39	75.0	481	MEC2_CABEL	Q27433 caenorhabdi
25	38	73.1	171	YSNB_BACSU	P94559 bacillus su
26	38	73.1	352	KE4_BRARE	Q9pub8 brachydanio
27	38	73.1	479	PRL2_ARATH	Q39190 arabidopsis
28	37	71.2	190	SLYD_HAEIN	P44830 haemophilus
29	37	71.2	191	HUPE_RHILV	P27650 rhizobium l
30	37	71.2	196	SLYD_ECOLI	P30856 escherichia
31	37	71.2	286	CH36_DROME	P07182 drosophila
32	36	69.2	97	NB2M_HUMAN	Q43676 homo sapien
33	36	69.2	242	PR38_YEAST	Q00723 saccharomyc

34	36	69.2	428	1	ITF2_RAT	Q62655 rattus norv
35	36	69.2	508	1	CPT7_HUMAN	P05093 homo sapien
36	36	69.2	642	1	ITF2_CANFA	P15881 canis famil
37	36	69.2	667	1	ITF2_HUMAN	P15884 homo sapien
38	36	69.2	670	1	ITF2_MOUSE	Q60722 mus musculu
39	35	67.3	112	1	SI09_MOUSE	P31725 mus musculu
40	35	67.3	118	1	SI09_RABIT	P50117 oryctolagus
41	35	67.3	187	1	UREE_ACTPL	O54422 actinobacil
42	35	67.3	235	1	Y057_CAEEL	Q09463 caenorhabdi
43	35	67.3	335	1	HYPB_RHOCA	P26410 rhodobacter
44	35	67.3	336	1	FILA_MOUSE	P11088 mus musculu
45	35	67.3	359	1	ZNT2_RAT	Q62941 rattus norv

ALIGNMENTS

RESULT 1						
MSA2_PLAFC						
ID	MSA2_PLAFC	STANDARD;	PRT;	262	AA.	
AC	Q99317;					
DT	01-OCT-1996 (Rel. 34, Created)					
DT	01-OCT-1996 (Rel. 34, Last sequence update)					
DT	01-OCT-1996 (Rel. 34, Last annotation update)					
DE	MEROZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2) (ALLELIC FORM 1).					
GN	MSA2.					
OS	Plasmodium falciparum (isolate Camp / Malaysia).					
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.					
OX	NCBI_taxid=5835;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE-91218803; PubMed-2090943;					
RA	Thomas A.W., Carr D.A., Carter J.M., Lyon J.A.;					
RT	"Sequence comparison of allelic forms of the Plasmodium falciparum					
RT	merozoite surface antigen MSA2."					
RL	Mol. Biochem. Parasitol. 43:211-220(1990).					
CC	!- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE					
CC	ERYTHROCYTE.					
CC	!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR					
CC	(POTENTIAL).					
CC	!- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/					
CC	or send an email to license@isb-sib.ch).					
CC	EMBL: M60186; AAA29687.1;					
DR	InterPro: IPR001136; MSA_2.					
DR	Pfam: PF00985; MSA_2; 1.					
KW	Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;					
KW	GPI-anchor; Merozoite.					
FT	SIGNAL 1 20					POTENTIAL.
FT	CHAIN 21 238					MEROZOITE SURFACE ANTIGEN 2.
FT	PROPEP 239 262					HYDROPHOBIC, REMOVED DURING MATURATION
FT						(BY SIMILARITY).
FT						POLYMORPHIC REGION.
FT	DOMAIN 44 188					POLY-THR.
FT	DOMAIN 91 98					N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 22 22					N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 36 36					N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 139 139					N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 211 211					N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 235 235					N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 236 236					N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE 262 AA; 27374 MW; 72E0B2A315E9D154 CRC64;					

Query Match 100.0%; Score 52; DB 1; Length 262;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 GQGHMHG 8
Db      196 GQGHMHG 203

RESULT 2
MSA2_PLAFF STANDARD; PRT; 264 AA.
AC      P19599;
AC      01-FEB-1991 (Rel. 17, Created)
DT      01-FEB-1991 (Rel. 17, Last sequence update)
DT      01-OCT-1996 (Rel. 34, Last annotation update)
DE      MEROZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2) (MEROZOITE 45 KDA
DE      SURFACE ANTIGEN) (AGS13).
MSA2
GN      Plasmodium falciparum (isolate FC27 / Papua New Guinea).
OS      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OC      NCBI_TaxID=5837;
RN      [1]
RX      MEDLINE=88276924; PubMed=3293051;
RA      Smythe J.A., Coppel R.L., Brown G.V., Ramasamy R., Kemp D.J.,
RA      Anders R.F.;
RT      Identification of two integral membrane proteins of Plasmodium
RT      falciparum.
RT      MEDLINE=88276924; PubMed=3293051;
RA      Smythe J.A., Coppel R.L., Brown G.V., Ramasamy R., Kemp D.J.,
RA      Anders R.F.;
RT      Identification of two integral membrane proteins of Plasmodium
RT      falciparum.
RL      Proc. Natl. Acad. Sci. U.S.A. 85:5195-5199(1988).
CC      -1- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE
CC      ERYTHROCYTE.
CC      -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
CC      (POTENTIAL).
CC      -1- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
CC      -1- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; J03828; AAA29627.1; ALT_SEQ.
DR      EMBL; A12418; CAA01031.1; -.
DR      PIR; A31818; A31818.
DR      InterPro; IPR001136; MSA_2.
DR      Pfam; PF00985; MSA_2; 1.
DR      Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
KW      GPI-anchor; Merozoite.
FT      SIGNAL 1 20
FT      CHAIN 21 240
FT      PROPEP 241 264
FT      -----
FT      DOMAIN 44 190
FT      DOMAIN 60 123
FT      REPEAT 60 91
FT      REPEAT 92 123
FT      CARBOHYD 22 22
FT      CARBOHYD 36 36
FT      CARBOHYD 213 213
FT      CARBOHYD 238 238
FT      SEQUENCE 264 AA; 27890 MW; 56DEF3B8357FEF3D CRC64;
SQ
Query Match 100.0%; Score 52; DB 1; Length 264;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GQGHMHG 8
Db      198 GQGHMHG 205

RESULT 3

```

DE SURFACE ANTIGEN).
GN MSA2.
OS Plasmodium falciparum (isolate 307).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90205972; PubMed=2181307;
RA Smythe J.A., Peterson M.G., Coppel R.L., Saul A.J., Kemp D.J.,
Anders R.F.;
RT "Structural diversity in the 45-kilodalton merozoite surface antigen
of Plasmodium falciparum";
RL Mol. Biochem. Parasitol. 39:227-234(1990).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE
ERYTHROCYTE.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
(POTENTIAL).
CC -1- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M28891; AAA29686.1; -.
DR InterPro: IPR001136; MSA_2.
DR Pfam: PF00985; MSA_2; 1.
KW Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
KW GPI-anchor; Merozoite.
FT SIGNAL 1 20
FT CHAIN 21 248
FT PROPEP 249 272
FT DOMAIN 44 198
FT DOMAIN 95 108
FT CARBOHYD 22 22
FT CARBOHYD 36 36
FT CARBOHYD 149 149
FT CARBOHYD 221 221
FT CARBOHYD 245 245
FT CARBOHYD 246 246
FT SEQUENCE 272 AA; 27971 MW; 9D9CF223BF2B483D CRC64;
Query Match 100.0%; Score 52; DB 1; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GQGHMHG 8
DB 206 GQGHMHG 213
RESULT 5
ID MSA2_PLAF6 STANDARD; PRT; 274 AA.
AC P50497;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MEROZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2).
GN MSA2.
OS Plasmodium falciparum (isolate kf1916).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57269;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92178286; PubMed=1542312;
RA Marshall V.M., Coppel R.L., Anders R.F., Kemp D.J.;
RT "Two novel alleles within subfamilies of the merozoite surface

RT antigen 2 (MSA-2) of Plasmodium falciparum";
RL Mol. Biochem. Parasitol. 50:181-184(1992).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE
ERYTHROCYTE.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
(POTENTIAL).
CC -1- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M73810; AAA29698.1; -.
DR InterPro: IPR001136; MSA_2.
DR Pfam: PF00985; MSA_2; 1.
KW Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
KW GPI-anchor; Merozoite.
FT SIGNAL 1 20
FT CHAIN 21 250
FT PROPEP 251 274
FT DOMAIN 44 200
FT DOMAIN 97 110
FT CARBOHYD 22 22
FT CARBOHYD 36 36
FT CARBOHYD 151 151
FT CARBOHYD 223 223
FT CARBOHYD 248 248
FT SEQUENCE 274 AA; 28367 MW; CEA832D766F743A2 CRC64;
Query Match 100.0%; Score 52; DB 1; Length 274;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GQGHMHG 8
DB 208 GQGHMHG 215
RESULT 6
ID MSA2_PLAF8 STANDARD; PRT; 276 AA.
AC Q99320;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MEROZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2) (ALLELIC FORM 4).
GN MSA2.
OS Plasmodium falciparum (isolate 7G8).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57266;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91218803; PubMed=2090943;
RA Thomas A.W., Carr D.A., Carter J.M., Lyon J.A.;
RT "Sequence comparison of allelic forms of the Plasmodium falciparum
merozoite surface antigen MSA2";
RL Mol. Biochem. Parasitol. 43:211-220(1990).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE
ERYTHROCYTE.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
(POTENTIAL).
CC -1- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)

CC -----
CC EMBL; M60190; AAA29690.1; -
CC InterPro; IPR001136; MSA_2;
CC Pfam; PF00985; MSA_2; 1.
CC Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
CC GPI-anchor; Merozoite.
CC SIGNAL 1 20 POTENTIAL.
CC CHAIN 21 252 MEROZOITE SURFACE ANTIGEN 2.
CC PROPEP 253 276 HYDROPHOBIC, REMOVED DURING MATURATION
CC (BY SIMILARITY).
CC FT DOMAIN 44 202 POLYMORPHIC REGION.
CC FT DOMAIN 105 112 POLY-THR.
CC FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 225 225 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 276 AA; 28172 MW; 85FA62A70400DEB2 CRC64;
SQ

Query Match 100.0%; Score 52; DB 1; Length 276;

Best Local Similarity 100.0%; Pred. No. 0.028;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8

DB 210 GQGHMHG 217

RESULT 7

MSA2_PLAFK
ID MSA2_PLAFK STANDARD; PRT; 280 AA.
AC Q03643;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MEROZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2).
GN MSA2.
OS Plasmodium falciparum (isolate K1 / Thailand).
OS Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5839;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91156685; PubMed=2000383;
RA Smythe J.A., Coppel R.L., Day K.P., Martin R.K., Oduola A.M.J.,
RA Kemp D.J., Anders R.F.;
RT "Structural diversity in the plasmodium falciparum merozoite surface
RT antigen 2.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:1751-1755(1991).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE
CC ERYTHROCYTE
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
CC (POTENTIAL).
CC -!- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; M59766; AAA29693.1; -
CC InterPro; IPR001136; MSA_2;
CC Pfam; PF00985; MSA_2; 1.
CC Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
CC GPI-anchor; Merozoite.
CC SIGNAL 1 20 POTENTIAL.
CC CHAIN 21 256 MEROZOITE SURFACE ANTIGEN 2.
CC

FT PROPEP 257 280 HYDROPHOBIC, REMOVED DURING MATURATION
FT DOMAIN 44 206 (BY SIMILARITY).
FT DOMAIN 160 163 POLYMORPHIC REGION.
FT CARBOHYD 22 22 POLY-SER.
FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 229 229 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 280 AA; 29542 MW; 9395D8AFB6B473EF CRC64;

Query Match 100.0%; Score 52; DB 1; Length 280;

Best Local Similarity 100.0%; Pred. No. 0.029;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8

DB 214 GQGHMHG 221

RESULT 8

MSA2_PLAFH
ID MSA2_PLAFH STANDARD; PRT; 281 AA.
AC Q99319;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MEROZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2) (ALLELIC FORM 3).
GN MSA2.
OS Plasmodium falciparum (isolate thtn / Thailand).
OS Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=70151;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91218803; PubMed=2090943;
RA Thomas A.W., Carr D.A., Carter J.M., Lyon J.A.;
RT "Sequence comparison of allelic forms of the Plasmodium falciparum
RT merozoite surface antigen MSA2.";
RL Mol. Biochem. Parasitol. 43:211-220(1990).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE
CC ERYTHROCYTE.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
CC (POTENTIAL).
CC -!- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; M60189; AAA29689.1; -
CC InterPro; IPR001136; MSA_2;
CC Pfam; PF00985; MSA_2; 1.
CC Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
CC GPI-anchor; Merozoite.
CC SIGNAL 1 20 POTENTIAL.
CC CHAIN 21 257 MEROZOITE SURFACE ANTIGEN 2.
CC PROPEP 258 281 HYDROPHOBIC, REMOVED DURING MATURATION
CC (BY SIMILARITY).
CC FT DOMAIN 44 207 POLYMORPHIC REGION.
CC FT DOMAIN 111 118 POLY-THR.
CC FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 281 AA; 28892 MW; 50598AA42D64CCBC CRC64;

Query Match 100.0%; Score 52; DB 1; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8
| | | | | | | |
Db 215 GQGHMHG 222

RESULT 9
MSA2_PLAF1 STANDARD; PRT; 286 AA.
AC P50496;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MEROZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2).
GN MSA2.
OS Plasmodium falciparum (isolate 311).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=57265;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92178286; PubMed=1542312;
RA Marshall V.M., Coppel R.L., Anders R.F., Kemp D.J.;
RT "Two novel alleles within subfamilies of the merozoite surface antigen 2 (MSA-2) of Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 50:181-184(1992).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE ERYTHROCYTE.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (POTENTIAL).
CC -!- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL; M73809; AAA29697.1; -
DR InterPro; IPR001136; MSA_2.
DR Pfam; PF00985; MSA_2; 1.
KW Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
FT SIGNAL 1 20
FT CHAIN 21 262
FT PROPEP 263 286
FT MEROZOITE SURFACE ANTIGEN 2.
FT HYDROPHOBIC, REMOVED DURING MATURATION
FT (BY SIMILARITY).
FT POLY-THR.
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 286 AA; 28844 MW; D1F4947CE68D5805 CRC64;

Query Match 100.0%; Score 52; DB 1; Length 286;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8
| | | | | | | |
Db 220 GQGHMHG 227

RESULT 10
MSA2_PLAFG STANDARD; PRT; 287 AA.
AC P19260;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MEROZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2) (ALLELIC FORM 2) (MEMBRANE PROTEIN PF7).
GN MSA2.
OS Plasmodium falciparum (isolate PCR-3 / Gambia).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5838;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90349616; PubMed=1696728;
RA Elliott J.F., Albrecht G.R., Gilladoga A., Handunnetti S.M., Needuaye J., Lallinger G., Minjas J.N., Howard R.J.;
RT "Genes for Plasmodium falciparum surface antigens cloned by expression in COS cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:6363-6367(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91117264; PubMed=1990294;
RA Fenton B., Clark J.T., Khan C.M.A., Robinson J.V., Walliker D., Ridley R., Scaife J.G., McBride J.S.;
RT "Structural and antigenic polymorphism of the 35- to 48-kilodalton merozoite surface antigen (MSA-2) of the malaria parasite Plasmodium falciparum.";
RL Mol. Cell. Biol. 11:963-971(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91218803; PubMed=2090943;
RA Thomas A.W., Carr D.A., Carter J.M., Lyon J.A.;
RT "Sequence comparison of allelic forms of the Plasmodium falciparum merozoite surface antigen MSA2.";
RL Mol. Biochem. Parasitol. 43:211-220(1990).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE ERYTHROCYTE.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (POTENTIAL).
CC -!- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL; M28890; AAA29650.1; -
DR EMBL; X53832; CAA37829.1; -
DR EMBL; M60188; AAA29688.1; -
DR PIR; A36018; A36018.
DR PIR; B39615; B39615.
DR InterPro; IPR001136; MSA_2.
DR Pfam; PF00985; MSA_2; 1.
KW Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
FT SIGNAL 1 20
FT CHAIN 21 263
FT PROPEP 264 287
FT MEROZOITE SURFACE ANTIGEN 2.
FT HYDROPHOBIC, REMOVED DURING MATURATION
FT (BY SIMILARITY).
FT POLYMORPHIC REGION.
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 287 AA; 28555 MW; 3968B90DAA917AF8 CRC64;

MSA2_PLAFZ		MSA2_PLAFZ	STANDARD;	PRT;	300 AA.
ID	AC	Q03645;			
DT	01-OCT-1996	(Rel. 34, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	01-OCT-1996	(Rel. 34, Last annotation update)			
DE	MEROZOITE SURFACE ANTIGEN 2 PRECURSOR (MSA-2).				

OC	Eukaryota; Alveolata; Apicomplexa, Haemosporida; Plasmodium.
NCBI_TaxID=70154;	
[1]	
RN	SEQUENCE FROM N.A.
RP	MEDLINE=9115685; PubMed=2000383;
RX	MDLINE=9115685; Day K.P., Martin R.K., Oduola A.M.J.,
RA	Smythe J.A., Coppel R.L., Kemp D.J., Anders R.F.;
RA	"Structural diversity in the Plasmodium falciparum merozoite surface
RT	antigen 2".
RT	Proc. Natl. Acad. Sci. U.S.A. 88:1751-1755(1991).
RL	-!- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE
CC	ERYTHROCYTE.
CC	-!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
CC	(POTENTIAL).
CC	-!- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; M59768; AAA29696.1; -
DR	PIR; A39112; A39112.
DR	InterPro; IPR001136; MSA_2.
DR	Pfam; PF00985; MSA_2; 1.
KW	Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
KW	GPI-anchor; Merozoite.
FT	SIGNAL 1 20
FT	CHAIN 21 276
FT	PROPEP 277 300
FT	DOMAIN 44 226
FT	DOMAIN 129 136
FT	CARBOHYD 22 22
FT	CARBOHYD 36 36
FT	CARBOHYD 177 177
FT	CARBOHYD 249 249
FT	CARBOHYD 273 273
FT	CARBOHYD 274 274
FT	SEQUENCE 300 AA; 30101 MW; E41161077AA10D CRC64;
SQ	
Query Match	100.0%; Score 52; DB 1; Length 300;
Best Local Similarity	100.0%; Pred. No. 0.031;
Matches 8; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 GQGHMHG 8
Db	234 GQGHMHG 241

RESULT	13		
MSA2_PLAF9			
ID	MSA2_PLAF9	STANDARD;	PRT; 302 AA.
AC	Q03994;		
DT	01-OCT-1996	(Rel. 34, Created)	
DT	01-OCT-1996	(Rel. 34, Last sequence update)	
DT	01-OCT-1996	(Rel. 34, Last annotation update)	

```

RA  Kemp D.J., Anders R.F.:
RT  "Structural diversity in the Plasmodium falciparum merozoite surface
RL  antigen 2.";
RL  Proc. Natl. Acad. Sci. U.S.A. 88:1751-1755(1991).
CC  -1- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE
CC  ERYTHROCYTE.
CC  -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
CC  (POTENTIAL).
CC  -1- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M59765; AAA29691.1; -.
DR  PIR; B39112; B39112.
DR  InterPro; IPR001136; MSA_2.
DR  Pfam; PF00985; MSA_2; 1.
KW  Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
KW  GPI-anchor; Merozoite.
FT  SIGNAL 1 20
FT  CHAIN 21 323
FT  PROPEP 324 347
FT  DOMAIN 44 273
FT  DOMAIN 177 184
FT  CARBOHYD 22 22
FT  CARBOHYD 36 36
FT  CARBOHYD 224 224
FT  CARBOHYD 296 296
FT  CARBOHYD 320 320
FT  CARBOHYD 321 321
SQ  SEQUENCE 347 AA; 33786 MW; ABCF24BB560BF537 CRC64;
Query Match 100.0%; Score 52; DB 1; Length 347;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GQGHMHG 8
DB 281 GQGHMHG 288
RESULT 15
CSUP_DROME ID CSUP_DROME STANDARD; PRT; 449 AA.
AC Q3V344;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CATECHOLAMINES UP PROTEIN.
GN CATSUP OR CG10449.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
ON [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99403013; PubMed=10471719;
RA Stathakis D.G., Burton D.Y., McIvor W.E., Krishnakumar S.,
RT Wright T.R., O'Donnell J.M.;
RT "The catecholamines up (Catsup) protein of Drosophila melanogaster
RT functions as a negative regulator of tyrosine hydroxylase activity.";
RL Genes 153:361-382(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.

```

Search completed: January 29, 2002, 11:13:45
Job time: 821 sec

```
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: NEGATIVELY REGULATES TYROSINE HYDROXYLASE ACTIVITY.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -!- SIMILARITY: BELONGS TO THE KE4/CATSUP FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF216384; AAF37226.1; -
DR EMBL; AE003661; AAF53744.1; -
DR FlyBase; FBgn002022; Catsup.
DR InterPro; IPR002395; Kininogen.
KW Transmembrane.
FT TRANSMEM 19 39 POTENTIAL.
FT TRANSMEM 135 155 POTENTIAL.
FT TRANSMEM 167 187 POTENTIAL.
FT TRANSMEM 222 242 POTENTIAL.
FT TRANSMEM 371 391 POTENTIAL.
FT TRANSMEM 395 415 POTENTIAL.
FT TRANSMEM 316 316 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 449 AA; 48658 MW; F7111A254C07AB4C CRC64;
SQ SEQUENCE
```

Query Match 78.8%; Score 41; DB 1; Length 449;
Best Local Similarity 75.08; Pred No 3.6;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GOGHMHG 8
I:|||||
Db 195 GEHGHG 202

6 3

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:12:13 ; Search time 285.36 Seconds
(without alignments)
4.101 Million cell updates/sec

Title: US-09-763-397A-15
Perfect score: 52
Sequence: 1 GQGHMHG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_17.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	129	5	O15688
2	52	100.0	137	5	O15689
3	52	100.0	147	5	O15805
4	52	100.0	152	5	O15683
5	52	100.0	157	5	Q9XZ38
6	52	100.0	161	5	Q9XZM2
7	52	100.0	163	5	Q9XZM4
8	52	100.0	164	5	O15690
9	52	100.0	167	5	O15686
10	52	100.0	171	5	O15685
11	52	100.0	174	5	Q9U0C4
12	52	100.0	176	5	Q9U0B0
13	52	100.0	176	5	Q9U0A7
14	52	100.0	176	5	Q9BJS4
15	52	100.0	176	5	Q9BJS3
16	52	100.0	178	5	Q9U0A1
17	52	100.0	179	5	Q9XZM3
18	52	100.0	179	5	Q9XZ40
19	52	100.0	180	5	Q25945

20	52	100.0	180	5	Q9TXC4	Q9txc4 plasmodium
21	52	100.0	183	5	Q9U0B6	Q9u0b6 plasmodium
22	52	100.0	183	5	Q9U0B5	Q9u0b5 plasmodium
23	52	100.0	184	5	Q9U0B9	Q9u0b9 plasmodium
24	52	100.0	184	5	Q9U0B8	Q9u0b8 plasmodium
25	52	100.0	184	5	Q9TY34	Q9ty34 plasmodium
26	52	100.0	186	5	Q9UB27	Q9ub27 plasmodium
27	52	100.0	187	5	Q9XZ39	Q9xz39 plasmodium
28	52	100.0	188	5	Q9U0A9	Q9u0a9 plasmodium
29	52	100.0	188	5	Q9U0A8	Q9u0a8 plasmodium
30	52	100.0	191	5	Q9U0C1	Q9u0c1 plasmodium
31	52	100.0	194	5	O15687	O15687 plasmodium
32	52	100.0	196	5	Q9U0B7	Q9u0b7 plasmodium
33	52	100.0	196	5	Q9GQY7	Q9ggy7 plasmodium
34	52	100.0	199	5	O15684	O15684 plasmodium
35	52	100.0	201	5	O15691	O15691 plasmodium
36	52	100.0	201	5	O9BJN4	O9bjn4 plasmodium
37	52	100.0	202	5	O15806	O15806 plasmodium
38	52	100.0	203	5	Q9U0C0	Q9u0c0 plasmodium
39	52	100.0	204	5	Q9U0B4	Q9u0b4 plasmodium
40	52	100.0	204	5	Q9U0A6	Q9u0a6 plasmodium
41	52	100.0	204	5	Q9U0A5	Q9u0a5 plasmodium
42	52	100.0	204	5	Q9U0A4	Q9u0a4 plasmodium
43	52	100.0	204	5	Q9U0A3	Q9u0a3 plasmodium
44	52	100.0	205	5	Q25944	Q25944 plasmodium
45	52	100.0	205	5	Q9TXC5	Q9txc5 plasmodium

ALIGNMENTS

RESULT 1	
O15688	
ID O15688	PRELIMINARY; PRT; 129 AA.
AC O15688	
DT 01-JAN-1998 (TREMBLrel. 05, Created)	
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)	
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	
DE MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).	
GN MSP2.	
OS Plasmodium falciparum.	
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
OX NCBI_TaxID=5833;	
RN [1]	
RP SEQUENCE FROM N.A.	
RA MEDLINE=98156776; PubMed=9497062;	
RA Irion A., Beck H.-P., Felger I.;	
RT "New repeat unit and hot spot of recombination in FC27-type alleles of the gene coding for Plasmodium falciparum merozoite surface protein 2.";	
RL Mol. Biochem. Parasitol. 90:367-370(1997).	
DR EMBL; AF010460; AAC38840.1; -	
DR InterPro; IPR001136; MSA_2.	
DR Pfam; PF00985; MSA_2; 1.	
KW Merozoite.	
FT NON_TER 1	
FT NON_TER 129 129	
SQ SEQUENCE 129 AA; 13114 MW; 741D7F5C2EB91A2E CRC64;	
Query Match 100.0%; Score 52; DB 5; Length 129;	
Best Local Similarity 100.0%; Pred. No. 0.09;	
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1 GQGHMHG 8	
Db 121 GQGHMHG 128	
RESULT 2	
O15689	
ID O15689	PRELIMINARY; PRT; 137 AA.
AC O15689;	

Mon Feb 4 15:23:38 2002

DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).
 GN MSP2.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98156776; PubMed=9497062;
 RA Irion A., Beck H.-P., Felger I.;
 RT "New repeat unit and hot spot of recombination in FC27-type alleles of
 the gene coding for Plasmodium falciparum merozoite surface protein
 2.";
 RL Mol. Biochem. Parasitol. 90:367-370(1997).
 DR EMBL; AF010455; AAC38835.1; -;
 DR InterPro; IPR001136; MSA_2.
 DR Pfam; PF00985; MSA_2; 1.
 KW Merozoite.
 FT NON_TER 1 152
 FT NON_TER 152 152
 SQ SEQUENCE 152 AA; 15268 MW; 0FE3718127D038A6 CRC64;

Query Match 100.0%; Score 52; DB 5; Length 152;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8
 |||||
 DB 144 GQGHMHG 151

RESULT 5
 Q9XZ38 PRELIMINARY; PRT; 157 AA.
 ID Q9XZ38
 AC Q9XZ38;
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).
 GN MSP2.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RHO212;
 RA Hosse R.J., Meisinger K., Engelbrecht F.;
 RT "Plasmodium falciparum msp2 allelic variants.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF132219; AAD26418.1; -;
 DR InterPro; IPR001136; MSA_2.
 DR Pfam; PF00985; MSA_2; 1.
 KW Merozoite.
 FT NON_TER 1 157
 FT NON_TER 157 157
 SQ SEQUENCE 157 AA; 15839 MW; 528D977160C2B00D CRC64;

Query Match 100.0%; Score 52; DB 5; Length 157;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8
 |||||
 DB 149 GQGHMHG 156

RESULT 6
 Q9XZM2 PRELIMINARY; PRT; 161 AA.
 ID Q9XZM2
 AC Q9XZM2;
 DT 01-NOV-1999 (TReMBLrel. 12, Created)

DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).
 GN MSP2.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98156776; PubMed=9497062;
 RA Irion A., Beck H.-P., Felger I.;
 RT "New repeat unit and hot spot of recombination in FC27-type alleles of
 the gene coding for Plasmodium falciparum merozoite surface protein
 2.";
 RL Mol. Biochem. Parasitol. 90:367-370(1997).
 DR EMBL; AF010461; AAC38841.1; -;
 DR InterPro; IPR001136; MSA_2.
 DR Pfam; PF00985; MSA_2; 1.
 KW Merozoite.
 FT NON_TER 1 137
 FT NON_TER 137 137
 SQ SEQUENCE 137 AA; 14053 MW; ED878BA4ECAABBS5D CRC64;

Query Match 100.0%; Score 52; DB 5; Length 137;
 Best Local Similarity 100.0%; Pred. No. 0.096;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8
 |||||
 DB 129 GQGHMHG 136

RESULT 3
 O15805 PRELIMINARY; PRT; 147 AA.
 ID O15805
 AC O15805;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Felger I., Irion A., Steiger S., Beck H.-P.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF033859; AAB87520.1; -;
 DR InterPro; IPR001136; MSA_2.
 DR Pfam; PF00985; MSA_2; 1.
 KW Merozoite.
 FT NON_TER 1 147
 FT NON_TER 147 147
 SQ SEQUENCE 147 AA; 14876 MW; 301B68169C222225 CRC64;

Query Match 100.0%; Score 52; DB 5; Length 147;
 Best Local Similarity 100.0%; Pred. No. 0.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8
 |||||
 DB 139 GQGHMHG 146

RESULT 4
 O15683 PRELIMINARY; PRT; 152 AA.
 ID O15683
 AC O15683;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DS MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).
 GN MSP2.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RHO210;
 RA Hosse R.J., Meisinger K., Engelbrecht F.;
 RL "Plasmodium falciparum msp2 allelic variants."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF148222; AAD37405.1;
 DR InterPro; IPR001136; MSA_2.
 DR Pfam; PF00985; MSA_2; 1.
 KW Merozoite.
 FT NON_TER 1 1
 FT NON_TER 161 161
 SQ SEQUENCE 161 AA; 15977 MW; B59CF23520F9BC02 CRC64;

Query Match 100.0%; Score 52; DB 5; Length 161;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8
 |||||
 Db 153 GQGHMHG 160

RESULT 7
 Q9XZM4 PRELIMINARY; PRT; 163 AA.
 ID Q9XZM4;
 AC Q9XZM4;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).
 GN MSP2.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AGU87;
 RA Hosse R.J., Meisinger K., Engelbrecht F.;
 RL "Plasmodium falciparum msp2 allelic variants."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF148224; AAD37407.1;
 DR InterPro; IPR001136; MSA_2.
 DR Pfam; PF00985; MSA_2; 1.
 KW Merozoite.
 FT NON_TER 1 1
 FT NON_TER 163 163
 SQ SEQUENCE 163 AA; 15612 MW; 1108FB27FAEB9677 CRC64;

Query Match 100.0%; Score 52; DB 5; Length 163;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8
 |||||
 Db 155 GQGHMHG 162

RESULT 8
 O15690 PRELIMINARY; PRT; 164 AA.
 ID O15690;
 AC O15690;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).
 GN MSP2.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98156776; PubMed=9497062;
 RA Irion A., Beck H.-P., Felger I.;
 RL "New repeat unit and hot spot of recombination in FC27-type alleles of
 RL the gene coding for Plasmodium falciparum merozoite surface protein
 RL 2."
 RL Mol. Biochem. Parasitol. 90:367-370(1997).
 DR EMBL; AF010452; AAC38842.1;
 DR InterPro; IPR001136; MSA_2.
 DR Pfam; PF00985; MSA_2; 1.
 KW Merozoite.
 FT NON_TER 1 1
 FT NON_TER 164 164
 SQ SEQUENCE 164 AA; 16686 MW; 5F5F49988D76C3CB CRC64;

Query Match 100.0%; Score 52; DB 5; Length 164;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8
 |||||
 Db 156 GQGHMHG 163

RESULT 9
 O15686 PRELIMINARY; PRT; 167 AA.
 ID O15686;
 AC O15686;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).
 GN MSP2.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98156776; PubMed=9497062;
 RA Irion A., Beck H.-P., Felger I.;
 RL "New repeat unit and hot spot of recombination in FC27-type alleles of
 RL the gene coding for Plasmodium falciparum merozoite surface protein
 RL 2."
 RL Mol. Biochem. Parasitol. 90:367-370(1997).
 DR EMBL; AF010458; AAC38838.1;
 DR InterPro; IPR001136; MSA_2.
 DR Pfam; PF00985; MSA_2; 1.
 KW Merozoite.
 FT NON_TER 1 1
 FT NON_TER 167 167
 SQ SEQUENCE 167 AA; 16774 MW; AF1419C6E8AC54E8 CRC64;

Query Match 100.0%; Score 52; DB 5; Length 167;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8
 |||||
 Db 159 GQGHMHG 166

RESULT 10
 O15685 PRELIMINARY; PRT; 171 AA.
 ID O15685

```

AC Q15685;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 2 (FRAGMENT).
GN MSP2.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98156776; PubMed=9497062;
RA Irion A., Beck H.-P., Felger I.;
RT "New repeat unit and hot spot of recombination in FC27-type alleles of
RT the gene coding for Plasmodium falciparum merozoite surface protein
RT 2."
RL Mol. Biochem. Parasitol. 90:367-370(1997).
DR EMBL; AF010457; AAC38837.1; -.
DR InterPro; IPR001136; MSA_2.
DR Pfam; PF00985; MSA_2; 1.
KW Merozoite.
KW NON_TER 1 1
FT NON_TER 171 171
FT SEQUENCE 171 AA; 16219 MW; 672EA3FC326A8F37 CRC64;

Query Match 100.0%; Score 52; DB 5; Length 171;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8
|||
DB 163 GQGHMHG 170

RESULT 11
Q9U0C4 PRELIMINARY; PRT; 174 AA.
AC Q9U0C4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MEROZOITE SURFACE ANTIGEN 2 (FRAGMENT).
GN MSP-2.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=V368;
RA Dobano C., Taylor R.R., McBride J.S.;
RT "Analysis of the structural diversity of Plasmodium falciparum
RT merozoite surface protein 2 (MSP-2) in Gambia."
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U91652; AAF21475.1; -.
DR InterPro; IPR001136; MSA_2.
DR Pfam; PF00985; MSA_2; 1.
KW Merozoite.
KW NON_TER 1 1
FT NON_TER 174 174
FT SEQUENCE 174 AA; 17582 MW; 7FA69C70996F8BED CRC64;
SQ

```

Q9U0B0	PRELIMINARY;	PRT;	176 AA.
ID	Q9U0B0		
AC	Q9U0B0;		
DT	01-MAY-2000 (TrEMBLrel. 13, Created)		
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)		
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)		
DE	MEROZOITE SURFACE ANTIGEN 2 (FRAGMENT).		
GN	MSP-2.		
OS	Plasmodium falciparum.		
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI_TaxID=5833;		
ON	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-V390;		
RA	Dobano C., Taylor R.R., McBride J.S.;		
RT	"Analysis of the structural diversity of Plasmodium falciparum		
RT	merozoite surface protein 2 (MSP-2) in Gambia.";		
RL	Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.		
DR	EWBL: U91666; AAF21489.1; -.		
DR	InterPro: IPR001136; MSA_2.		
KW	Pfam: PF00985; MSA_2; 1.		
KW	Merozoite.		
FT	1		
FT	NON_TER	176	
SQ	SEQUENCE	176 AA; 18083 MW; DB89779EB5B03DEB CRC64;	
Query Match 100.0%; Score 52; DB 5; Length 176;			
Best Local Similarity 100.0%; Pred. No. 0.12;			
Matches 8; Conservative 0; Mismatches 0; Indels 0;			
Qy	1 GQGHMHG 8		
Db	169 GQGHMHG 176		
RESULT	13		
Q9U0A7	PRELIMINARY;	PRT;	176 AA.
ID	Q9U0A7		
AC	Q9U0A7;		
DT	01-MAY-2000 (TrEMBLrel. 13, Created)		
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)		
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)		
DE	MEROZOITE SURFACE ANTIGEN 2 (FRAGMENT).		
GN	MSP-2.		
OS	Plasmodium falciparum.		
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI_TaxID=5833;		
ON	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-V342;		
RA	Dobano C., Taylor R.R., McBride J.S.;		
RT	"Analysis of the structural diversity of Plasmodium falciparum		
RT	merozoite surface protein 2 (MSP-2) in Gambia.";		
RL	Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.		
DR	EWBL: U91669; AAF21492.1; -.		
DR	InterPro: IPR001136; MSA_2.		
DR	Pfam: PF00985; MSA_2; 1.		
KW	Merozoite.		
FT	1		
FT	NON_TER	176	
SQ	SEQUENCE	176 AA; 18142 MW; A13140FCE312A0D2 CRC64;	

RESULT

RESULT 14

Q9BJS4
ID Q9BJS4 PRELIMINARY; PRT; 176 AA.
AC Q9BJS4;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MERZOITE SURFACE PROTEIN 2 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Clifton N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329539; AAK19358.1; -
FT NON_TER 1
FT NON_TER 176
SQ SEQUENCE 176 AA; 18256 MW; 3FC1A48233972B63 CRC64;

Query Match 100.0%; Score 52; DB 5; Length 176;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8

Db 159 GQGHMHG 166

RESULT 15

Q9BJS3
ID Q9BJS3 PRELIMINARY; PRT; 176 AA.
AC Q9BJS3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MERZOITE SURFACE PROTEIN 2 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Eisen D.P., Clifton N., Baddeley A., Eri R., Saul A.;
RT "Antigenic drift and immune selection acting on merozoite surface
RT protein 1-19 and merozoite surface protein 2 in independent field
RT isolates of Plasmodium falciparum."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329540; AAK19359.1; -
FT NON_TER 1
FT NON_TER 176
SQ SEQUENCE 176 AA; 18256 MW; 3FC1A48233972B63 CRC64;

Query Match 100.0%; Score 52; DB 5; Length 176;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8

Db 159 GQGHMHG 166

Search completed: January 29, 2002, 11:12:13
Job time: 769 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:21:47 ; Search time 310.82 Seconds
(without alignments)
4.290 Million cell updates/sec

Title: US-09-763-397a-16

Perfect score: 100

Sequence: 1 DGNCEIDPHVNEFSAIDL 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- A_Geneseq_1101.*
- 1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT.*
 - 2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT.*
 - 3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT.*
 - 4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT.*
 - 5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT.*
 - 6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT.*
 - 7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT.*
 - 8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT.*
 - 9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT.*
 - 10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT.*
 - 11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT.*
 - 12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT.*
 - 13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT.*
 - 14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT.*
 - 15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT.*
 - 16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT.*
 - 17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT.*
 - 18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT.*
 - 19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT.*
 - 20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT.*
 - 21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT.*
 - 22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	100.0	18	21	AA770292 Plasmodium falciparum
2	100	100.0	350	21	AA770278 Recombinant vaccinia
3	100	100.0	622	10	AA916332 Rhostry membrane a
4	95	95.0	622	13	AA927532 Plasmodium falciparum
5	95	95.0	622	16	AA968840 Plasmodium falciparum
6	66	66.0	563	12	AA910551 Plasmodium knowlesi
7	65	65.0	489	12	AA910935 Plasmodium vivax 6
8	47.5	47.5	135	21	AA933060 Zea mays protein f
9	47.5	47.5	137	21	AAG28496 Zea mays protein f
10	47.5	47.5	146	21	AAG33054 Zea mays protein f
11	47.5	47.5	160	21	AAG35922 Zea mays protein f

12	47.5	47.5	160	21	AAG40638 Zea mays protein f
13	47.5	47.5	163	21	AAG33058 Zea mays protein f
14	47.5	47.5	173	21	AAG33053 Zea mays protein f
15	47.5	47.5	192	21	AAG40637 Zea mays protein f
16	44	44.0	1804	19	AAW64178 Lettuce resistance
17	43.5	43.5	115	21	AAG34638 Arabidopsis thaliana
18	43.5	43.5	138	21	AAG22810 Arabidopsis thaliana
19	43.5	43.5	139	21	AAG51650 Arabidopsis thaliana
20	43.5	43.5	142	21	AAG34637 Arabidopsis thaliana
21	43.5	43.5	158	21	AAG08047 Arabidopsis thaliana
22	43.5	43.5	158	21	AAG31610 Arabidopsis thaliana
23	43.5	43.5	158	22	AAB66879 Arabidopsis eif-5A
24	43.5	43.5	159	21	AAG11630 Arabidopsis thaliana
25	43.5	43.5	159	21	AAG51644 Arabidopsis thaliana
26	43.5	43.5	160	22	AAB66877 Tomato eif-5A, Ly
27	43.5	43.5	164	21	AAG22809 Arabidopsis thaliana
28	43.5	43.5	164	21	AAG51649 Arabidopsis thaliana
29	43.5	43.5	200	21	AAG34426 Arabidopsis thaliana
30	43.5	43.5	202	21	AAG11629 Arabidopsis thaliana
31	43.5	43.5	202	21	AAG51643 Arabidopsis thaliana
32	43.5	43.5	252	21	AAG34425 Arabidopsis thaliana
33	43.5	43.5	294	21	AAG34424 Arabidopsis thaliana
34	43	43.0	135	21	AAG26588 Arabidopsis thaliana
35	43	43.0	138	21	AAG26587 Arabidopsis thaliana
36	43	43.0	145	18	AAW15106 Myxoma virus MA51
37	43	43.0	182	18	AAW15100 Myxoma virus MA51
38	43	43.0	242	21	AAG44660 Zea mays protein f
39	43	43.0	283	21	AAG44659 Zea mays protein f
40	43	43.0	304	21	AAG44658 Zea mays protein f
41	43	43.0	1073	7	AAP60569 Sequence of the RI
42	42.5	42.5	133	21	AAG27163 Zea mays protein f
43	42.5	42.5	160	21	AAG27161 Zea mays protein f
44	42	42.0	412	22	AA01038 CFE 41 protein seq
45	42	42.0	1009	21	AA79169 Pneumocystis carin

ALIGNMENTS

RESULT 1	
AA770292	AA770292 standard; peptide; 18 AA.
ID	AA770292 standard; peptide; 18 AA.
XX	
AC	AA770292;
XX	
DT	06-JUN-2000 (first entry)
XX	
DE	Plasmodium falciparum AMA-1 antigenic epitope, P600.
XX	
KW	Recombinant protein; CDC/NIAIDVAC-1; multivalent; malaria; vaccine;
KW	T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
KW	Circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
KW	liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
KW	apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
KW	EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
KW	Pfg27; antiparasitic; prevention; anti-CDC/NIAIDVAC-1 antibody.
OS	Plasmodium falciparum.
XX	
PN	WO200011179-A1.
XX	
PD	02-MAR-2000.
XX	
PF	19-AUG-1999; 99WO-US18869.
XX	
PR	21-AUG-1998; 98US-0097703.
XX	
PA	(NATM-) NAT INST IMMUNOLOGY.
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX	
PI	Lal AA, Shi YP, Hasnain SE;
XX	
DR	WPI; 2000-237654/20.

XX Novel recombinant protein as vaccine for treating malarial infection
PT comprises antigenic peptides obtained from different stages of
PT plasmodium falciparum life cycle -
XX
PS Claim 2; Page 16; 52pp; English.
XX
CC The present sequence is the antigenic epitope P600, derived from
CC apical membrane antigen-1 (AMA-1) of the asexual blood stage of
CC Plasmodium falciparum. It is used in the construction of recombinant
CC protein CDC/NIIMALVAC-1, which is a multivalent, multistage malarial
CC vaccine. The recombinant protein comprises, melittin signal peptide,
CC (His)6 tag, T-cell epitope from tetanus toxoid and 21 antigenic epitopes
CC from circumsporozoite protein (CSP), sporozoite surface protein-2
CC (SSP-2), liver stage antigen-1 (LSA-1), merozoite surface protein-1
CC (MSP-1), MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding
CC antigen-175 (EBA-175), rhothry associated protein-1 (RAP-1) and gamete
CC specific antigen, Pf27. These epitopes were obtained at different stages
CC of the life cycle of P. falciparum. CDC/NIIMALVAC-1 vaccine has
CC antiparasitic activity and can be used for treatment and prevention of
CC malarial infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for
XX detecting P. falciparum in biological samples.
XX Sequence 18 AA;
SQ
Query Match 100.0%; Score 100; DB 21; Length 18;
Best Local Similarity 100.0%; Pred. NO. 3.7e-10; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 0;
QY 1 DGNCEIDPHVNEFSALDL 18
DB 1 DGNCEIDPHVNEFSALDL 18
RESULT 2
AAY70278
ID AAY70278 standard; Protein; 350 AA.
XX
AC AAY70278;
XX
DT 06-JUN-2000 (first entry)
XX
DE Recombinant vaccine CDC/NIIMALVAC-1.
XX
KW Recombinant protein: CDC/NIIMALVAC-1; multivalent; malaria; vaccine;
KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
KW EBA-175; rhothry associated protein-1; RAP-1; gamete specific antigen;
KW Pf27; antiparasitic; prevention; anti-CDC/NIIMALVAC-1 antibody;
KW honey bee.
XX
OS Chimeric - Apis sp.
OS Chimeric - Clostridium tetani.
OS Chimeric - Plasmodium falciparum.
XX
XX Key Location/Qualifiers
FT Peptide 1..22
FT /label= Melittin_signal_peptide
FT /note= "Derived from Honey bee"
FT Protein 23..350
FT /label= Mature_CDC/NIIMALVAC-1
FT /note= "Recombinant multivalent malarial vaccine"
XX
XX WO200011179-A1.
XX
XX 02-MAR-2000.
XX
XX 19-AUG-1999; 99WO-US18869.
XX
XX 21-AUG-1998; 98US-0097703.

XX (NAIM-) NAT INST IMMUNOLOGY.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Lal AA, Shi YP, Hasnain SE;
PI
XX WPI; 2000-237654/20.
DR N-PSDB; AAZ51336.
XX
XX Novel recombinant protein as vaccine for treating malarial infection
PT comprises antigenic peptides obtained from different stages of
PT plasmodium falciparum life cycle -
XX
PS Claim 3; Page 43-44; 52pp; English.
XX
CC The present sequence is that of recombinant protein CDC/NIIMALVAC-1,
CC which is a multivalent, multistage malarial vaccine. The recombinant
CC protein comprises, melittin signal peptide, (His)6 tag, T-cell epitope
CC from tetanus toxoid and 21 antigenic epitopes from circumsporozoite
CC protein (CSP), sporozoite surface protein-2 (SSP-2), liver stage
CC antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1), MSP-2, apical
CC membrane antigen-1 (AMA-1), erythrocyte binding antigen-175 (EBA-175),
CC rhothry associated protein-1 (RAP-1) and gamete specific antigen, Pf27.
CC These epitopes were obtained at different stages of the life cycle of
CC Plasmodium falciparum. CDC/NIIMALVAC-1 vaccine has antiparasitic
CC activity and can be used for treatment and prevention of malarial
CC infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for detecting
XX P. falciparum in biological samples.
XX Sequence 350 AA;
SQ
Query Match 100.0%; Score 100; DB 21; Length 350;
Best Local Similarity 100.0%; Pred. NO. 1.1e-08; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 0;
QY 1 DGNCEIDPHVNEFSALDL 18
DB 140 dgnceidphvnefsaldl 157
RESULT 3
AAP91632
ID AAP91632 standard; protein; 622 AA.
XX
AC AAP91632;
XX
DT 25-JAN-1990 (first entry)
XX
DE Rhothry membrane antigen-1.
XX
KW Malaria; rhothry membrane antigen-1; antibodies.
XX
OS Plasmodium falciparum D10.
XX
XX WO8907645-A.
XX
XX 24-AUG-1989.
XX
XX 10-FEB-1989; 89WO-AU000056.
XX
XX 12-FEB-1988; 88AU-0006743.
XX
XX (SARA) SARAMANE PTY LTD.
XX
XX Peterson MG, Crewther PE, Smythe JA, Marshall VM, Silva A;
PI
XX WPI; 1989-263714/36.
DR N-PSDB; AAN90703.
XX
XX Rhothry membrane antigen of Plasmodium falciparum
PT - used for producing antibodies and in immunisation,
PT diagnostic and treatment methods for malaria.
PT

XX Claim 1; Fig 3; 46pp; English.
 PS
 CC RMA-1 can generate an immune response to malaria, and antibodies which
 CC can inhibit growth of the parasite. RMA-1 initially has mol. wt. 80 kD.
 XX
 SQ Sequence 622 AA;

Query Match 100.0%; Score 100; DB 10; Length 622;
 Best Local Similarity 100.0%; Pred. No. 2.1e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DGNCEIDPHVNEFSAIDL 18
 |||||

Db 317 dgcncediphvnefsaidl 334

RESULT 4
 AAR27532
 ID AAR27532 standard; Protein: 622 AA.
 AC AAR27532;
 XX
 DT 08-MAR-1993 (first entry)
 XX
 DE Plasmodium falciparum AMA1 antigen.
 XX
 KW Recombinant poxvirus; antimalarial vaccine; malaria; immunise;
 KW immunogen.
 XX
 OS Plasmodium falciparum.
 XX
 PN WO9216616-A.
 XX
 PD 01-OCT-1992.
 XX
 PF 19-MAR-1992; 92WO-US02207.
 XX
 PR 20-MAR-1991; 91US-0672183.
 PR 18-MAR-1992; 92US-0852305.
 XX
 PA (VIRO-) VIROGENETICS CORP.
 XX
 PI De Taisne C, Paoletti E, Tine JA;
 XX
 DR WPI; 1992-349203/42.
 DR N-PSDB; AAQ29189.
 XX
 PT Recombinant poxvirus - contg. Plasmodium DNA, useful as
 PT antimalarial vaccine
 XX
 PS Example 4; Fig 5; 74pp; English.

XX This sequence is the plasmodium falciparum AMA1 antigen.
 CC cDNA encoding it was cloned into vaccinia donor plasmids
 CC before being inserted into the vaccinia virus to be used in a
 CC vaccine to stimulate an antimalarial immunological response, or for
 CC in vitro prodn. of gene prods. for use as immunogens. As plasmodium
 CC genes are conserved among P. falciparum strains, they are widely
 CC effective in a vaccine.
 XX
 SQ Sequence 622 AA;

Query Match 95.0%; Score 95; DB 13; Length 622;
 Best Local Similarity 94.4%; Pred. No. 1.4e-07;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DGNCEIDPHVNEFSAIDL 18
 |||||

Db 317 dgcncediphvnefsaidl 334

RESULT 5
 AAR68840
 ID AAR68840 standard; Protein: 622 AA.
 XX
 AC AAR68840;
 XX
 DT 24-AUG-1995 (first entry)
 XX
 DE Plasmodium falciparum AMA-1 gene protein.
 XX
 KW Plasmodium falciparum AMA-1 gene; recombinant poxvirus;
 KW multicomponent multistage malarial vaccines; immunogens;
 KW malaria diagnosis.
 XX
 OS Plasmodium falciparum (3D7).
 XX
 PN WO9428930-A.
 XX
 PD 22-DEC-1994.
 XX
 PF 10-JUN-1994; 94WO-US06652.
 XX
 PR 11-JUN-1993; 93US-0075783.
 PR 09-JUN-1994; 94US-0257073.
 XX
 PA (VIRO-) VIROGENETICS CORP.
 XX
 PI De Taisne C, Paoletti E, Tine JA;
 XX
 DR WPI; 1995-036113/05.
 DR N-PSDB; AAQ80910.
 XX
 PT Recombinant poxvirus contg. Plasmodium DNA in non-essential
 PT region - useful in vaccines against malaria and for prodn. of
 PT Plasmodium immunogens
 XX
 PS Claim 3; Fig 5; 183pp; English.
 XX
 CC AA080910 encodes AAR68840 the P. falciparum AMA-1 gene product. New
 CC recombinant poxviruses containing either the SERA, ABRA, Pfhspp70,
 CC AMA-1, Pfs25, Pfs16, CSP, PfSSP2, LSA-1, LSA-1 repeatless, MSA-1,
 CC MSA-1 (N-terminal p83 or C-terminal gp42) genes, or a combination
 CC of these in non-essential regions of their genomes are claimed.
 CC These poxviruses (pref. with a virulence reducing genomic
 CC deletion or disruption) can be used as vaccines against malaria
 CC and for the prodn. of plasmodium immunogens. These viruses
 CC provide multicomponent, multistage vaccines due to their expression
 CC of sporozite, liver stage, blood stage and sexual stage proteins.
 XX
 SQ Sequence 622 AA;

Query Match 95.0%; Score 95; DB 16; Length 622;
 Best Local Similarity 94.4%; Pred. No. 1.4e-07;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DGNCEIDPHVNEFSAIDL 18
 |||||

Db 317 dgcncediphvnefsaidl 334

RESULT 6
 AAR10551
 ID AAR10551 standard; Protein: 563 AA.
 XX
 AC AAR10551;
 XX
 DT 15-APR-1991 (first entry)
 XX
 DE Plasmodium knowlesi 66kD merozoite antigen protein.
 XX
 KW Plasmodium merozoite antigens; malaria vaccine.

XX OS Plasmodium knowlesi.
 XX PN US7483516-A.
 XX PD 15-JAN-1991.
 XX PF 22-FEB-1990; 90US-0483516.
 XX PR 22-FEB-1990; 90US-0483516.
 XX PA (USSH) NAT INST OF HEALTH.
 XX PI Waters AP, McCutchan TF;
 XX DR WPI; 1991-044381/06.
 XX DR N-PSDB; AAQ10418.
 XX PT Plasmodium merozoite antigen proteins and DNA sequences - useful
 XX in vaccine prodn., anti-malarial drug design, and in diagnostics
 XX PS Disclosure; fig 1; 3lpp; English.
 XX CC This protein is encoded by a recombinant DNA molecule comprising
 XX the complete P.knowlesi merozoite antigen gene and e.g. a pUC19
 XX or vaccinia virus vector sequence and regulatory elements. This
 XX allows efficient expression of the antigen on transformation of
 XX host cells. It is useful in an anti-malarial vaccine for con-
 XX ferring immunity against the merozoite form of the parasite and
 XX preventing infection of uninfected red blood cells.
 XX CC See also AAQ10869.
 XX SQ Sequence 563 AA;
 XX
 XX Query Match 66.0%; Score 56; DB 12; Length 563;
 XX Best Local Similarity 66.7%; Pred. No. 0.0091;
 XX Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 XX
 XX QY 1 DGNCEIDIPHVNEFSAIDL 18
 XX |||||:||||:| | |
 XX Db 262 dgnceeiptynearsl 279
 XX
 XX RESULT 7
 XX AAR10935
 XX ID AAR10935 standard; Protein; 489 AA.
 XX AC AAR10935;
 XX DT 15-APR-1991 (first entry)
 XX DE Plasmodium vivax 66kD merozoite antigen protein partial sequence.
 XX KW Plasmodium merozoite antigens; malaria vaccine.
 XX OS Plasmodium vivax.
 XX PN US7483516-A.
 XX PD 15-JAN-1991.
 XX PF 22-FEB-1990; 90US-0483516.
 XX PR 22-FEB-1990; 90US-0483516.
 XX PA (USSH) NAT INST OF HEALTH.
 XX PI Waters AP, McCutchan TF;
 XX DR WPI; 1991-044381/06.
 XX DR N-PSDB; AAQ10869.
 XX
 XX Query Match 65.0%; Score 65; DB 12; Length 489;
 XX Best Local Similarity 66.7%; Pred. No. 0.011;
 XX Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 XX
 XX QY 1 DGNCEIDIPHVNEFSAIDL 18
 XX |||||:||||:| | |
 XX Db 188 dgnceeiptyveagdl 205
 XX
 XX RESULT 8
 XX AAG33060
 XX ID AAG33060 standard; Protein; 135 AA.
 XX AC AAG33060;
 XX DT 17-OCT-2000 (first entry)
 XX DE Zea mays protein fragment SEQ ID NO: 40000.
 XX KW protein identification; signal transduction pathway; metabolic pathway;
 XX hybridization assay; genetic mapping; gene expression control; promoter;
 XX termination sequence; corn.
 XX OS Zea mays subsp. mays.
 XX PN EF1033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-0301439.
 XX PR 25-FEB-1999; 99US-0121825.
 XX PR 05-MAR-1999; 99US-0123180.
 XX PR 09-MAR-1999; 99US-0123548.
 XX PR 23-MAR-1999; 99US-0125788.
 XX PR 25-MAR-1999; 99US-0126264.
 XX PR 29-MAR-1999; 99US-0126785.
 XX PR 01-APR-1999; 99US-0127462.
 XX PR 06-APR-1999; 99US-0128234.
 XX PR 08-APR-1999; 99US-0128714.
 XX PR 16-APR-1999; 99US-0129845.
 XX PR 19-APR-1999; 99US-0130077.
 XX PR 21-APR-1999; 99US-0130449.
 XX PR 23-APR-1999; 99US-0130510.
 XX PR 23-APR-1999; 99US-0130891.
 XX PR 28-APR-1999; 99US-0131449.
 XX PR 30-APR-1999; 99US-0132048.
 XX PR 30-APR-1999; 99US-0132407.
 XX PR 04-MAY-1999; 99US-0132484.
 XX PR 05-MAY-1999; 99US-0132485.
 XX PR 06-MAY-1999; 99US-0132486.
 XX PR 07-MAY-1999; 99US-0132487.
 XX PR 11-MAY-1999; 99US-0132863.
 XX PR 14-MAY-1999; 99US-0134256.
 XX PR 14-MAY-1999; 99US-0134218.
 XX PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145210.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149802.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

```
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 47.5%; Score 47.5; DB 21; Length 135;
Best Local Similarity 61.1%; Pred. No. 2.2;
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 2;

Oy 3 NCEDIPHN--EFSAILD 18
Db 81 nc-diphvnrteyqlidi 97

RESULT 9
AAG28496
ID AAG28496 standard; Protein: 137 AA.
XX
AC AAG28496;
XX
DT 17-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 33735.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
PD
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131149.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.

25-MAY-1999; 99US-0136021.
27-MAY-1999; 99US-0136392.
28-MAY-1999; 99US-0136782.
01-JUN-1999; 99US-0137222.
03-JUN-1999; 99US-0137528.
04-JUN-1999; 99US-0137502.
07-JUN-1999; 99US-0137724.
08-JUN-1999; 99US-0138094.
10-JUN-1999; 99US-0138540.
10-JUN-1999; 99US-0138847.
14-JUN-1999; 99US-0139119.
16-JUN-1999; 99US-0139452.
16-JUN-1999; 99US-0139453.
17-JUN-1999; 99US-0139492.
18-JUN-1999; 99US-0139454.
18-JUN-1999; 99US-0139455.
18-JUN-1999; 99US-0139456.
18-JUN-1999; 99US-0139457.
18-JUN-1999; 99US-0139458.
18-JUN-1999; 99US-0139459.
18-JUN-1999; 99US-0139460.
18-JUN-1999; 99US-0139461.
18-JUN-1999; 99US-0139462.
18-JUN-1999; 99US-0139463.
18-JUN-1999; 99US-0139750.
18-JUN-1999; 99US-0139763.
21-JUN-1999; 99US-0139817.
22-JUN-1999; 99US-0139899.
23-JUN-1999; 99US-0140353.
23-JUN-1999; 99US-0140354.
24-JUN-1999; 99US-0140595.
28-JUN-1999; 99US-0140823.
29-JUN-1999; 99US-0140991.
30-JUN-1999; 99US-0141287.
01-JUL-1999; 99US-0141842.
01-JUL-1999; 99US-0142154.
02-JUL-1999; 99US-0142055.
06-JUL-1999; 99US-0142390.
08-JUL-1999; 99US-0142803.
09-JUL-1999; 99US-0142920.
12-JUL-1999; 99US-0142977.
13-JUL-1999; 99US-0143542.
14-JUL-1999; 99US-0143624.
15-JUL-1999; 99US-0144005.
16-JUL-1999; 99US-0144085.
16-JUL-1999; 99US-0144086.
19-JUL-1999; 99US-0144325.
19-JUL-1999; 99US-0144331.
19-JUL-1999; 99US-0144332.
19-JUL-1999; 99US-0144333.
19-JUL-1999; 99US-0144334.
19-JUL-1999; 99US-0144335.
19-JUL-1999; 99US-0144352.
20-JUL-1999; 99US-0144632.
20-JUL-1999; 99US-0144884.
21-JUL-1999; 99US-0144814.
21-JUL-1999; 99US-0145086.
21-JUL-1999; 99US-0145088.
22-JUL-1999; 99US-0145085.
22-JUL-1999; 99US-0145087.
22-JUL-1999; 99US-0145089.
22-JUL-1999; 99US-0145192.
23-JUL-1999; 99US-0145145.
23-JUL-1999; 99US-0145218.
23-JUL-1999; 99US-0145224.
26-JUL-1999; 99US-0145276.
27-JUL-1999; 99US-0145913.
27-JUL-1999; 99US-0145918.
27-JUL-1999; 99US-0145919.
28-JUL-1999; 99US-0145951.
02-AUG-1999; 99US-0146386.
02-AUG-1999; 99US-0146388.
02-AUG-1999; 99US-0146389.
```

```
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 47.5%; Score 47.5; DB 21; Length 137;
Best Local Similarity 61.1%; Pred. No. 2.3;
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 2;

QY 3 NCEDIPHVN--EFSAILD 18
Db 112 nc-diphvnrteyqlidi 128
|| ||||| |::
|| ||||| |::

RESULT 10
AAG33054
ID AAG33054 standard; Protein; 146 AA.
XX
AC AAG33054;
XX
DT 17-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 39992.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
25-FEB-2000; 2000EP-0301439.
XX
25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
```

us-09-763-397a-16.rag

Mon Feb 4 15:23:38 2002

PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144329.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145273.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 47.5%; DB 21; Length 146;
Best Local Similarity 61.1%;
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 2;

PR	16-AUG-1999;	99US-0149368.	ID	AAG40638 standard; Protein; 160 AA.
PR	17-AUG-1999;	99US-0149175.	XX	
PR	18-AUG-1999;	99US-0149426.	AC	AAG40638;
PR	20-AUG-1999;	99US-0149722.	XX	
PR	20-AUG-1999;	99US-0149723.	DT	18-OCT-2000 (first entry)
PR	20-AUG-1999;	99US-0149929.	XX	
PR	23-AUG-1999;	99US-0149902.	DE	Zea mays protein fragment SEQ ID NO: 50450.
PR	23-AUG-1999;	99US-0149930.	XX	
PR	25-AUG-1999;	99US-0150366.	KW	Protein identification; signal transduction pathway; metabolic pathway;
PR	26-AUG-1999;	99US-0150884.	KW	hybridisation assay; genetic mapping; gene expression control; promoter;
PR	27-AUG-1999;	99US-0151065.	KW	termination sequence; corn.
PR	27-AUG-1999;	99US-0151066.	XX	
PR	27-AUG-1999;	99US-0151080.	XX	Zea mays subsp. mays.
PR	30-AUG-1999;	99US-0151303.	XX	
PR	31-AUG-1999;	99US-0151438.	XX	EP1033405-A2.
PR	01-SEP-1999;	99US-0151930.	PN	
PR	07-SEP-1999;	99US-0152363.	PD	06-SEP-2000.
PR	10-SEP-1999;	99US-0153070.	XX	
PR	13-SEP-1999;	99US-0153758.	XX	25-FEB-2000; 2000EP-0301439.
PR	15-SEP-1999;	99US-0154018.	XX	
PR	16-SEP-1999;	99US-0154039.	PR	25-FEB-1999; 99US-0121825.
PR	20-SEP-1999;	99US-0154779.	PR	05-MAR-1999; 99US-0123180.
PR	22-SEP-1999;	99US-0155139.	PR	09-MAR-1999; 99US-0123548.
PR	23-SEP-1999;	99US-0155486.	PR	23-MAR-1999; 99US-0125788.
PR	24-SEP-1999;	99US-0155459.	PR	25-MAR-1999; 99US-0126264.
PR	28-SEP-1999;	99US-0156458.	PR	29-MAR-1999; 99US-0126785.
PR	29-SEP-1999;	99US-0156596.	PR	01-APR-1999; 99US-0127462.
PR	04-OCT-1999;	99US-0157117.	PR	06-APR-1999; 99US-0128234.
PR	05-OCT-1999;	99US-0157753.	PR	08-APR-1999; 99US-0128714.
PR	06-OCT-1999;	99US-0157865.	PR	16-APR-1999; 99US-0129845.
PR	07-OCT-1999;	99US-0158029.	PR	19-APR-1999; 99US-0130077.
PR	08-OCT-1999;	99US-0158232.	PR	21-APR-1999; 99US-0130449.
PR	12-OCT-1999;	99US-0158369.	PR	23-APR-1999; 99US-0130510.
PR	13-OCT-1999;	99US-0159293.	PR	23-APR-1999; 99US-0130891.
PR	13-OCT-1999;	99US-0159294.	PR	28-APR-1999; 99US-0131449.
PR	13-OCT-1999;	99US-0159295.	PR	30-APR-1999; 99US-0132048.
PR	14-OCT-1999;	99US-0159329.	PR	30-APR-1999; 99US-0132407.
PR	14-OCT-1999;	99US-0159330.	PR	04-MAY-1999; 99US-0132484.
PR	14-OCT-1999;	99US-0159331.	PR	05-MAY-1999; 99US-0132485.
PR	14-OCT-1999;	99US-0159637.	PR	06-MAY-1999; 99US-0132486.
PR	14-OCT-1999;	99US-0159638.	PR	08-MAY-1999; 99US-0132487.
PR	18-OCT-1999;	99US-0159584.	PR	07-MAY-1999; 99US-0132863.
PR	21-OCT-1999;	99US-0160741.	PR	11-MAY-1999; 99US-0134256.
PR	21-OCT-1999;	99US-0160767.	PR	14-MAY-1999; 99US-0134218.
PR	21-OCT-1999;	99US-0160768.	PR	14-MAY-1999; 99US-0134221.
PR	21-OCT-1999;	99US-0160770.	PR	14-MAY-1999; 99US-0134370.
PR	21-OCT-1999;	99US-0160814.	PR	18-MAY-1999; 99US-0134768.
PR	21-OCT-1999;	99US-0160815.	PR	19-MAY-1999; 99US-0134941.
PR	22-OCT-1999;	99US-0160980.	PR	20-MAY-1999; 99US-0135124.
PR	22-OCT-1999;	99US-0160981.	PR	21-MAY-1999; 99US-0135353.
PR	22-OCT-1999;	99US-0160989.	PR	24-MAY-1999; 99US-0135629.
PR	25-OCT-1999;	99US-0161404.	PR	25-MAY-1999; 99US-0136021.
PR	25-OCT-1999;	99US-0161405.	PR	27-MAY-1999; 99US-0136392.
PR	25-OCT-1999;	99US-0161406.	PR	28-MAY-1999; 99US-0136782.
PR	26-OCT-1999;	99US-0161359.	PR	01-JUN-1999; 99US-0137222.
PR	26-OCT-1999;	99US-0161360.	PR	03-JUN-1999; 99US-0137528.
PR	26-OCT-1999;	99US-0161361.	PR	04-JUN-1999; 99US-0137502.
PR	28-OCT-1999;	99US-0161920.	PR	07-JUN-1999; 99US-0137724.
PR	28-OCT-1999;	99US-0161992.	PR	08-JUN-1999; 99US-0138094.
PR	28-OCT-1999;	99US-0161993.	PR	10-JUN-1999; 99US-0138540.
PR	29-OCT-1999;	99US-0162142.	PR	10-JUN-1999; 99US-0138847.
PR			PR	10-JUN-1999; 99US-0139119.
PR			PR	14-JUN-1999; 99US-0139452.
PR			PR	16-JUN-1999; 99US-0139453.
PR			PR	16-JUN-1999; 99US-0139454.
PR			PR	17-JUN-1999; 99US-0139492.
PR			PR	18-JUN-1999; 99US-0139454.
PR			PR	18-JUN-1999; 99US-0139455.
PR			PR	18-JUN-1999; 99US-0139456.
PR			PR	18-JUN-1999; 99US-0139457.
PR			PR	18-JUN-1999; 99US-0139458.
PR			PR	18-JUN-1999; 99US-0139459.
PR			PR	18-JUN-1999; 99US-0139460.

Query Match 47.5%; Score 47.5; DB 21; Length 160;
Best Local Similarity 61.1%; Pred. No. 2.7;
Matches 11; Conservative 2; Mismatches 3; Indels 3; Gaps 2;

QY 3 NCEDIPHN--EFSIDL 18
||| ||||| |
Db 81 nc-diphvnrteyqlidi 97

RESULT 12
AAG40638

PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 47.5%; Score 47.5; DB 21; Length 160;
Best Local Similarity 61.1%; Pred. No. 2.7;
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 2;

Qy 3 NCEDIPHVN--EFSAILD 18
Db 81 nc-diphvnrteyqlidi 97

RESULT 13
AAG33058
ID AAG33058 standard; Protein; 163 AA.
XX
AC AAG33058;
XX
DT 17-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 39998.

XX	Protein identification; signal transduction pathway; metabolic pathway;	PR 23-JUN-1999;	99US-0140353;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	PR 23-JUN-1999;	99US-0140354;
KW	termination sequence; corn.	PR 24-JUN-1999;	99US-0140695;
XX		PR 28-JUN-1999;	99US-0140823.
XX		PR 29-JUN-1999;	99US-0140991.
OS		PR 30-JUN-1999;	99US-0141287.
XX	Zea mays subsp. mays.	PR 01-JUL-1999;	99US-0141842.
XX		PR 01-JUL-1999;	99US-0142154.
PN	EP1033405-A2.	PR 02-JUL-1999;	99US-0142055.
PD		PR 06-JUL-1999;	99US-0142390.
XX	06-SEP-2000.	PR 08-JUL-1999;	99US-0142803.
XX		PR 09-JUL-1999;	99US-0142920.
PF	25-FEB-2000; 2000EP-0301439.	PR 12-JUL-1999;	99US-0142977.
XX		PR 13-JUL-1999;	99US-0143542.
PR	25-FEB-1999;	PR 14-JUL-1999;	99US-0143624.
PR	05-MAR-1999;	PR 15-JUL-1999;	99US-0144005.
PR	09-MAR-1999;	PR 16-JUL-1999;	99US-0144085.
PR	23-MAR-1999;	PR 16-JUL-1999;	99US-0144086.
PR	25-MAR-1999;	PR 19-JUL-1999;	99US-0144325.
PR	29-MAR-1999;	PR 19-JUL-1999;	99US-0144331.
PR	01-APR-1999;	PR 19-JUL-1999;	99US-0144332.
PR	08-APR-1999;	PR 19-JUL-1999;	99US-0144333.
PR	06-APR-1999;	PR 19-JUL-1999;	99US-0144334.
PR	16-APR-1999;	PR 19-JUL-1999;	99US-0144335.
PR	19-APR-1999;	PR 20-JUL-1999;	99US-0144352.
PR	21-APR-1999;	PR 20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	PR 21-JUL-1999;	99US-0144884.
PR	28-APR-1999;	PR 21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	PR 21-JUL-1999;	99US-0145086.
PR	08-APR-1999;	PR 21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	PR 22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	PR 22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	PR 22-JUL-1999;	99US-0145089.
PR	06-MAY-1999;	PR 22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	PR 23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	PR 23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	PR 23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	PR 26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	PR 27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	PR 27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	PR 27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	PR 28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	PR 02-AUG-1999;	99US-0145951.
PR	21-MAY-1999;	PR 02-AUG-1999;	99US-0146386.
PR	24-MAY-1999;	PR 02-AUG-1999;	99US-0146388.
PR	25-MAY-1999;	PR 03-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	PR 03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	PR 04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	PR 04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	PR 05-AUG-1999;	99US-0147302.
PR	03-JUN-1999;	PR 05-AUG-1999;	99US-0147192.
PR	04-JUN-1999;	PR 06-AUG-1999;	99US-0147260.
PR	07-JUN-1999;	PR 06-AUG-1999;	99US-0147303.
PR	08-JUN-1999;	PR 06-AUG-1999;	99US-0147416.
PR	10-JUN-1999;	PR 09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	PR 09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	PR 10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	PR 11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	PR 12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	PR 13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	PR 13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	PR 16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	PR 17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	PR 18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	PR 20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	PR 20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	PR 20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR 23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	PR 23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	PR 25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	PR 26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	PR 27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	PR 27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	PR 27-AUG-1999;	99US-015108

PR	31-AUG-1999;	99US-0151438.	PN	EP1033405-A2.
PR	01-SEP-1999;	99US-0151930.	XX	
PR	07-SEP-1999;	99US-0152363.	PD	
PR	10-SEP-1999;	99US-0153070.	XX	06-SEP-2000.
PR	13-SEP-1999;	99US-0153758.	PF	
PR	15-SEP-1999;	99US-0154018.	XX	25-FEB-2000; 2000EP-0301439.
PR	16-SEP-1999;	99US-0154039.	PR	25-FEB-1999;
PR	20-SEP-1999;	99US-0154779.	PR	99US-0121825.
PR	22-SEP-1999;	99US-0155139.	PR	99US-0123180.
PR	23-SEP-1999;	99US-0155139.	PR	99US-0123548.
PR	24-SEP-1999;	99US-0155486.	PR	99US-0125788.
PR	28-SEP-1999;	99US-0155659.	PR	99US-0126264.
PR	29-SEP-1999;	99US-0156458.	PR	99US-0126785.
PR	29-SEP-1999;	99US-0156596.	PR	99US-0126785.
PR	04-OCT-1999;	99US-0157117.	PR	99US-0127462.
PR	05-OCT-1999;	99US-0157753.	PR	99US-0128234.
PR	06-OCT-1999;	99US-0157865.	PR	99US-0128714.
PR	07-OCT-1999;	99US-0158029.	PR	99US-0128845.
PR	08-OCT-1999;	99US-0158232.	PR	99US-0130077.
PR	12-OCT-1999;	99US-0158369.	PR	99US-0130449.
PR	13-OCT-1999;	99US-0159293.	PR	99US-0130510.
PR	13-OCT-1999;	99US-0159294.	PR	99US-0130891.
PR	13-OCT-1999;	99US-0159295.	PR	99US-0131449.
PR	14-OCT-1999;	99US-0159329.	PR	99US-0132048.
PR	14-OCT-1999;	99US-0159330.	PR	99US-0132407.
PR	14-OCT-1999;	99US-0159331.	PR	99US-0132484.
PR	14-OCT-1999;	99US-0159637.	PR	99US-0132485.
PR	14-OCT-1999;	99US-0159638.	PR	99US-0132486.
PR	18-OCT-1999;	99US-0159584.	PR	99US-0132487.
PR	21-OCT-1999;	99US-0160741.	PR	99US-0132863.
PR	21-OCT-1999;	99US-0160767.	PR	99US-0134256.
PR	21-OCT-1999;	99US-0160768.	PR	99US-0134218.
PR	21-OCT-1999;	99US-0160770.	PR	99US-0134219.
PR	21-OCT-1999;	99US-0160814.	PR	99US-0134221.
PR	21-OCT-1999;	99US-0160815.	PR	99US-0134370.
PR	22-OCT-1999;	99US-0160980.	PR	99US-0134768.
PR	22-OCT-1999;	99US-0160981.	PR	99US-0134941.
PR	22-OCT-1999;	99US-0160989.	PR	99US-0135124.
PR	25-OCT-1999;	99US-0161404.	PR	99US-0135353.
PR	25-OCT-1999;	99US-0161405.	PR	99US-0135629.
PR	25-OCT-1999;	99US-0161406.	PR	99US-0136021.
PR	26-OCT-1999;	99US-0161359.	PR	99US-0136392.
PR	26-OCT-1999;	99US-0161360.	PR	99US-0136782.
PR	26-OCT-1999;	99US-0161361.	PR	99US-0137222.
PR	28-OCT-1999;	99US-0161920.	PR	99US-0137528.
PR	28-OCT-1999;	99US-0161992.	PR	99US-0137502.
PR	28-OCT-1999;	99US-0161993.	PR	99US-0137724.
PR	29-OCT-1999;	99US-0162142.	PR	99US-0138094.
PR			PR	99US-0138540.
PR			PR	99US-0138847.
PR			PR	99US-0139119.
PR			PR	99US-0139452.
PR			PR	99US-0139453.
PR			PR	99US-0139492.
PR			PR	99US-0139454.
PR			PR	99US-0139455.
PR			PR	99US-0139456.
PR			PR	99US-0139457.
PR			PR	99US-0139458.
PR			PR	99US-0139459.
PR			PR	99US-0139460.
PR			PR	99US-0139461.
PR			PR	99US-0139462.
PR			PR	99US-0139463.
PR			PR	99US-0139750.
PR			PR	99US-0139763.
PR			PR	99US-0139817.
PR			PR	99US-0139899.
PR			PR	99US-0139899.
PR			PR	99US-0140353.
PR			PR	99US-0140354.
PR			PR	99US-0140695.
PR			PR	99US-0140823.
PR			PR	99US-0140991.
PR			PR	99US-0141287.
PR			PR	99US-0141842.

Query Match 47.5%; Score 47.5; DB 21; Length 163;
Best Local Similarity 61.1%; Pred. No. 2.8;
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 2;
OY 3 NCEDIPHVN--EFSIDL 18
DB 109 nc-diphvnrteyqlidi 125
|| ||||| | : ||:
nc-diphvnrteyqlidi 125
RESULT 14
AAG33053
ID AAG33053 standard; Protein: 173 AA.
XX AC AAG33053;
XX DT 17-OCT-2000 (first entry)
XX DE Zea mays protein fragment SEQ ID NO: 39991.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX OS Zea mays subsp. mays.
XX XX

PR 01-JUL-1999;	99US-0142154.	PR 20-SEP-1999;	99US-0154779.
PR 02-JUL-1999;	99US-0142055.	PR 22-SEP-1999;	99US-0155139.
PR 06-JUL-1999;	99US-0142390.	PR 23-SEP-1999;	99US-0155486.
PR 08-JUL-1999;	99US-0142803.	PR 24-SEP-1999;	99US-0155659.
PR 09-JUL-1999;	99US-0142920.	PR 28-SEP-1999;	99US-0156458.
PR 12-JUL-1999;	99US-0142977.	PR 29-SEP-1999;	99US-0156596.
PR 13-JUL-1999;	99US-0143542.	PR 04-OCT-1999;	99US-0157117.
PR 14-JUL-1999;	99US-0143624.	PR 05-OCT-1999;	99US-0157753.
PR 15-JUL-1999;	99US-0144005.	PR 06-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144085.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144086.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144325.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0158293.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144333.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144884.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0144814.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160767.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160815.
PR 22-JUL-1999;	99US-0145192.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160989.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0161404.
PR 26-JUL-1999;	99US-0145276.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145918.	PR 26-OCT-1999;	99US-0161359.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161360.
PR 28-JUL-1999;	99US-0145951.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146386.	PR 28-OCT-1999;	99US-0161920.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161992.
PR 02-AUG-1999;	99US-0146389.	PR 28-OCT-1999;	99US-0161993.
PR 03-AUG-1999;	99US-0147038.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147204.	PR 29-OCT-1999;	99US-0162142.
PR 04-AUG-1999;	99US-0147302.		
PR 05-AUG-1999;	99US-0147192.		
PR 05-AUG-1999;	99US-0147260.		
PR 06-AUG-1999;	99US-0147303.		
PR 06-AUG-1999;	99US-0147416.		
PR 09-AUG-1999;	99US-0147493.		
PR 09-AUG-1999;	99US-0147935.		
PR 10-AUG-1999;	99US-0148171.		
PR 11-AUG-1999;	99US-0148319.		
PR 12-AUG-1999;	99US-0148341.		
PR 13-AUG-1999;	99US-0148565.		
PR 13-AUG-1999;	99US-0148684.		
PR 16-AUG-1999;	99US-0149368.		
PR 17-AUG-1999;	99US-0149175.		
PR 18-AUG-1999;	99US-0149426.		
PR 20-AUG-1999;	99US-0149722.		
PR 20-AUG-1999;	99US-0149723.		
PR 20-AUG-1999;	99US-0149929.		
PR 23-AUG-1999;	99US-0149902.		
PR 23-AUG-1999;	99US-0149930.		
PR 25-AUG-1999;	99US-0150566.		
PR 26-AUG-1999;	99US-0150884.		
PR 27-AUG-1999;	99US-0151065.		
PR 27-AUG-1999;	99US-0151066.		
PR 27-AUG-1999;	99US-0151080.		
PR 30-AUG-1999;	99US-0151303.		
PR 31-AUG-1999;	99US-0151438.		
PR 01-SEP-1999;	99US-0151930.		
PR 07-SEP-1999;	99US-0152363.		
PR 10-SEP-1999;	99US-0153070.		
PR 13-SEP-1999;	99US-0153758.		
PR 15-SEP-1999;	99US-0154018.		
PR 16-SEP-1999;	99US-0154039.		
<hr/>			
Query Match 47.5%; Score 47.5; DB 21; Length 173;			
Best Local Similarity 61.1%; Pred. No. 2.9;			
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 2;			
QY 3 NCEDIPHVN--EFSADL 18			
:			
Db 108 nc-diphvnrteyqlidi 124			
<hr/>			
RESULT 15			
AAG40637			
ID AAG40637 standard; Protein; 192 AA.			
XX AC AAG40637;			
XX AC			
XX DT 18-OCT-2000 (first entry)			
XX DE			
XX DE Zea mays protein fragment SEQ ID NO: 50449.			
XX KW Protein identification; signal transduction pathway; metabolic pathway;			
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;			
XX KW termination sequence; corn.			
XX OS Zea mays subsp. mays.			
XX OS			
XX PN EP1033405-A2.			
XX PN			
XX PD 06-SEP-2000.			
XX PD			
XX XX 25-FEB-2000; 2000EP-0301439.			
XX XX			
XX PR 25-FEB-1999; 99US-0121825.			

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

```
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

```
Query Match 47.5%; Score 47.5; DB 21; Length 192;
Best Local Similarity 61.1%; Pred. No. 3.3;
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 2;
```

```
QY 3 NCEDIPHVN--EFSAILD 18
Db 113 nc-diphvnrteyqlidi 129
```

Search completed: January 29, 2002, 10:21:48
Job time: 424 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:24:07 ; Search time 133.18 seconds
(without alignments)
3.041 Million cell updates/sec

Title: US-09-763-397A-16
Perfect score: 100
Sequence: 1 DGNCEIPHNEFSATDL 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95	95.0	628	1	US-08-257-073-9
2	43	43.0	145	4	US-09-068-655-11
3	43	43.0	182	4	US-09-068-655-4
4	41	41.0	660	3	US-09-058-489-18
5	41	41.0	660	3	US-09-058-489-91
6	40	40.0	432	1	US-07-698-926A-2
7	40	40.0	926	1	US-08-159-340A-2
8	39	39.0	314	3	US-09-080-044-3
9	39	39.0	423	4	US-08-943-714-9
10	38.5	38.5	447	4	US-09-370-253-10
11	38.5	38.5	696	6	5262177-4
12	38.5	38.5	698	2	US-08-175-158A-2
13	38.5	38.5	1074	2	US-08-470-058-2
14	38.5	38.5	1074	3	US-09-037-188-2
15	38.5	38.5	1074	4	US-09-285-310-2
16	38.5	38.5	1410	2	US-08-470-058-4
17	38.5	38.5	1410	3	US-09-037-188-4
18	38.5	38.5	1410	4	US-09-285-310-4
19	38	38.0	135	3	US-08-950-720A-2
20	37.5	37.5	447	4	US-09-370-253-2
21	37	37.0	41	6	5512648-1
22	37	37.0	337	2	US-08-861-464-12
23	37	37.0	337	2	US-08-396-001-12
24	37	37.0	337	4	US-09-323-433A-12
25	37	37.0	483	2	US-08-416-870C-4
26	37	37.0	500	1	US-07-980-526-2
27	37	37.0	500	1	US-08-221-817-15

28	37	37.0	500	1	US-08-454-439-15	Sequence 15, Appl
29	37	37.0	500	5	PCT-US94-10487-15	Sequence 15, Appl
30	37	37.0	573	2	US-08-991-531-1	Sequence 1, Appl
31	37	37.0	573	2	US-08-991-531-10	Sequence 10, Appl
32	37	37.0	573	2	US-08-993-318A-10	Sequence 10, Appl
33	37	37.0	573	3	US-09-028-887-1	Sequence 1, Appl
34	37	37.0	573	4	US-09-399-886-10	Sequence 10, Appl
35	37	37.0	573	4	US-09-396-260-10	Sequence 10, Appl
36	37	37.0	573	4	US-09-518-901-1	Sequence 1, Appl
37	37	37.0	573	4	US-09-576-281-10	Sequence 10, Appl
38	37	37.0	620	1	US-08-706-037-27	Sequence 27, Appl
39	37	37.0	620	1	US-08-940-661A-2	Sequence 2, Appl
40	37	37.0	620	2	US-09-083-485-2	Sequence 2, Appl
41	37	37.0	620	2	US-09-005-397-27	Sequence 27, Appl
42	37	37.0	620	2	US-08-939-218A-2	Sequence 2, Appl
43	37	37.0	620	5	PCT-US95-06815-2	Sequence 2, Appl
44	37	37.0	1119	4	US-09-396-651B-2	Sequence 2, Appl
45	37	37.0	1290	1	US-08-470-350B-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-257-073-9
; Sequence 9, Application US/08257073
; Patent No. 5766597
; GENERAL INFORMATION:
; APPLICANT: Paolletti, Enzo
; APPLICANT: de Taisne, Charles
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue, 25th Floor
; CITY: New York
; STATE: New York
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,073
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/075,783
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/852,305
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,183
; FILING DATE: 20-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2570
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 628 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

```
; FRAGMENT TYPE: internal
US-08-257-073-9

Query Match          95.0%; Score 95; DB 1; Length 628;
Best Local Similarity 94.4%; Pred. No. 8.9e-08;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DGCNEDIPHVNEFSAIDL 18
Db 317 DGCNEDIPHVNEFPAIDL 334
|||||
RESULT 2
US-09-068-655-11
; Sequence 11, Application US/09068655A
; Patent No. 6136579
; GENERAL INFORMATION:
; APPLICANT: JACKSON, Ronald James
; TITLE OF INVENTION: METHOD OF PRODUCING ALPHA 2,3-SIALYLTRANSFERASE
; FILE REFERENCE: 026579-186
; CURRENT APPLICATION NUMBER: US/09/068,655A
; CURRENT FILING DATE: 1998-08-31
; EARLIER FILING DATE: 1995-11-15
; EARLIER FILING DATE: 1995-11-15
; EARLIER FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Myxoma virus
US-09-068-655-11

Query Match          43.0%; Score 43; DB 4; Length 145;
Best Local Similarity 53.3%; Pred. No. 5.8;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 DGCNEDIPHVNEFSA 15
Db 87 DGLCIEIKHVGRFGA 101
|||||
RESULT 3
US-09-068-655-4
; Sequence 4, Application US/09068655A
; Patent No. 6136579
; GENERAL INFORMATION:
; APPLICANT: JACKSON, Ronald James
; TITLE OF INVENTION: METHOD OF PRODUCING ALPHA 2,3-SIALYLTRANSFERASE
; FILE REFERENCE: 026579-186
; CURRENT APPLICATION NUMBER: US/09/068,655A
; CURRENT FILING DATE: 1998-08-31
; EARLIER FILING DATE: 1995-11-15
; EARLIER FILING DATE: 1995-11-15
; EARLIER FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Myxoma virus
US-09-068-655-4

Query Match          43.0%; Score 43; DB 4; Length 182;
Best Local Similarity 53.3%; Pred. No. 7.5;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 DGCNEDIPHVNEFSA 15
Db 87 DGLCIEIKHVGRFGA 101
|||||
RESULT 4
US-09-058-489-18
; Sequence 18, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of the Y Chromosome
; FILE REFERENCE: WHI97-08pA
; CURRENT APPLICATION NUMBER: US/09/058,489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041,877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Human
US-09-058-489-18

Query Match          41.0%; Score 41; DB 3; Length 660;
Best Local Similarity 50.0%; Pred. No. 69;
Matches 8; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

QY 3 NCEDIPHVNEFSAIDL 18
Db 172 NCP--PHIENFSDIM 185
|||||
RESULT 5
US-09-058-489-91
; Sequence 91, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of the Y Chromosome
; FILE REFERENCE: WHI97-08pA
; CURRENT APPLICATION NUMBER: US/09/058,489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041,877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 91
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Human
US-09-058-489-91

Query Match          41.0%; Score 41; DB 3; Length 660;
Best Local Similarity 50.0%; Pred. No. 69;
Matches 8; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

QY 3 NCEDIPHVNEFSAIDL 18
Db 172 NCP--PHIENFSDIM 185
|||||
RESULT 6
US-07-698-926A-2
; Sequence 2, Application US/07698926A
```


APPLICANT: Golightly, Elizabeth
 APPLICANT: Klotz, Alan
 APPLICANT: Mathisen, Thomas Erik
 APPLICANT: Dammann, Claus
 TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
 TITLE OF INVENTION: Encoding Same
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESS: NO. 6187578o No. 6187578disk of No. 6187578th America, Inc.
 STREET: 405 Lexington Avenue
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10174
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/943,714
 FILING DATE: 03-OCT-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Lambiris, Elias J
 REGISTRATION NUMBER: 33,728
 REFERENCE/DOCKET NUMBER: 4990.200-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-867-0123
 TELEFAX: 212-878-9655
 TELEX:
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 423 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-943-714-9

Query Match 39.0%; Score 39; DB 4; Length 423;
Best Local Similarity 53.8%; Pred. No. 89;
Matches 7; Conservative 2; Mismatches 4; Indels

Qy 2 GNCEDIPHVNEFS 14
| | | | | | | | | |
Db 6 GICETTPGVNOYS 18

```

RESULT 10
US-09-370-253-10
    : SEQUENCE 10, Application US/09370253
    : Patent No. 6165792
    : GENERAL INFORMATION:
    : APPLICANT: Allen, Stephen M.
    : APPLICANT: Sakai, Hajime
    : APPLICANT: Thorpe, Catherine J.
    : TITLE OF INVENTION: Amino Acid Transporters
    : FILE REFERENCE: BB-1200
    : CURRENT APPLICATION NUMBER: US/09/370,253
    : CURRENT FILING DATE: 1999-08-09
    : EARLIER APPLICATION NUMBER: 60/097,222
    : EARLIER FILING DATE: August 20, 1998
    : NUMBER OF SEQ ID NOS: 18
    : SOFTWARE: Microsoft Office 97
    : SEQ ID NO 10
    : LENGTH: 447
    : TYPE: PRT
    : ORGANISM: Triticum aestivum
US-09-370-253-10

```

Query Match
38.5%; Score 38.5; DB 4; Length 447;

```

Best Local Similarity 24.3%; Pred. No. 1.1e+02; Gaps 1;
Matches 9; Conservative 4; Mismatches 5; Indels 19; Gaps 1;

QY 1 DGCNEDI-----PHVNEFSAILD 18
   |||::|
   |||::|
Db 151 DGCKDKIKLFFIMFASVHVFLVSOLFNLNSISGVSL 187

RESULT 11
5262177-4
; Patent No. 5262177
; APPLICANT: BROWN, J OSEPH P.; ESTIN, CHARLES D.; PLOWMAN, GREGORY
; . D.; HELLSTROM, KARL E.; ROSE, TIMOTHY M.; HELLSTROM, INGEGERD;
; PURCHIO, ANTHONY F.; HU, SHIU-LOK; PENNATHUR, SRIDHAR
; TITLE OF INVENTION: RECOMBINANT VIRUSES ENCODING THE HUMAN
; MELANOMA-ASSOCIATED ANTIGEN
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/7230
; FILING DATE: 27-JAN-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 827,313
; FILING DATE: 07-FEB-1986
; SEQ ID NO:4:
; LENGTH: 696
5262177-4

Query Match 38.5%; Score 38.5; DB 6; Length 696;
Best Local Similarity 43.5%; Pred. No. 1.9e+02; Gaps 1;
Matches 10; Conservative 0; Mismatches 6; Indels 7; Gaps 1;

```

Query Match 38.5%; Score 38.5; DB 6; Length 696;
Best Local Similarity 43.5%; Pred. No. 1.9e+02;
Matches 10: Conservative 0; Mismatches 6; Indels

QY 3 NCEDIPVNEF-----SAIDL 18
||||| | | | |
Db 435 NCEDTPEAGYFAVAVVKKSASDL 457

RESULT 12
US-08-175-158A-2
; Sequence 2, Application US/08175158A
; Patent No. 5986067
; GENERAL INFORMATION:
; APPLICANT: FUNK, Walter D.
; APPLICANT: MACGILLIVRAY, Ross T.A.
; APPLICANT: MASON, Anne B.
; APPLICANT: WOODWORTH, Robert C.
; TITLE OF INVENTION: RECOMBINANT TRANSFERRINS, TRANSFERRIN HALP-
; TITLE OF INVENTION: MOLECULES AND MUTANTS THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/175,158A
; FILING DATE: 28-DEC-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/832,029
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: UVI-005Cp2
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 698 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-175-158A-2

Query Match 38.5%; Score 38.5; DB 2; Length 698;
Best Local Similarity 43.5%; Pred. No. 1.9e+02;
Matches 10; Conservative 0; Mismatches 6; Indels 7; Gaps 1;

QY 3 NCEDIPHVNEF-----SAIDL 18
||||| | | | | | | | | |
Db 436 NCEDTPEAGYFAVAVVKSASDL 458

RESULT 13
US-08-470-058-2

; Sequence 2, Application US/08470058
; Patent No. 5817789
; GENERAL INFORMATION:
; APPLICANT: Heartlein, Michael W.
; APPLICANT: Lemontt, Jeffrey F.
; TITLE OF INVENTION: Chimeric Proteins For Use in Transport
; TITLE OF INVENTION: Of a Selected Substance Into Cells
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470.058
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: TKT93-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1074 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-470-058-2

Query Match 38.5%; Score 38.5; DB 2; Length 1074;
Best Local Similarity 43.5%; Pred. No. 3.1e+02;
Matches 10; Conservative 0; Mismatches 6; Indels 7; Gaps 1;

QY 3 NCEDIPHVNEF-----SAIDL 18
||||| | | | | | | | | |
Db 812 NCEDTPEAGYFAVAVVKSASDL 834

RESULT 14

US-09-037-188-2
; Sequence 2, Application US/09037188
; Patent No. 6027921
; GENERAL INFORMATION:

; APPLICANT: Heartlein, Michael W.
; APPLICANT: Lemontt, Jeffrey F.
; APPLICANT: Concino, Michael F.

; TITLE OF INVENTION: CHIMERIC PROTEINS FOR USE
; TITLE OF INVENTION: IN TRANSPORT OF A SELECTED SUBSTANCE INTO CELLS
; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA

; COUNTRY: USA
; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows95

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/037,188

; FILING DATE: 02-MAR-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE: 06-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Fraser, Ph.D., J.D., Janis K.

; REGISTRATION NUMBER: 34,819

; REFERENCE/DOCKET NUMBER: 07236/009002

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1074 amino acids

; TYPE: amino acid

; TOPOLOGY: unknown

; MOLECULE TYPE: protein
US-09-037-188-2

Query Match 38.5%; Score 38.5; DB 3; Length 1074;
Best Local Similarity 43.5%; Pred. No. 3.1e+02;
Matches 10; Conservative 0; Mismatches 6; Indels 7; Gaps 1;

QY 3 NCEDIPHVNEF-----SAIDL 18
||||| | | | | | | | | |
Db 812 NCEDTPEAGYFAVAVVKSASDL 834

RESULT 15

US-09-285-310-2

; Sequence 2, Application US/09285310

; Patent No. 6262026

; GENERAL INFORMATION:

; APPLICANT: Heartlein, Michael W.

; APPLICANT: Lemontt, Jeffrey F.

; APPLICANT: Concino, Michael F.

; TITLE OF INVENTION: CHIMERIC PROTEINS FOR USE

; TITLE OF INVENTION: IN TRANSPORT OF A SELECTED SUBSTANCE INTO CELLS

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

```
;
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/285,310
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/037,188
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Ph.D., J.D., Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 07236/009002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1074 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
;
US-09-285-310-2

Query Match      38.5%; Score 38.5; DB 4; Length 1074;
Best Local Similarity 43.5%; Pred. No. 3.1e+02;
Matches 10; Conservative 0; Mismatches 6; Indels 7; Gaps 1;

Oy 3 NCEDIPHVNEF-----SAIDL 18
   ||||| | | |||
Db 812 NCEDTPEAGYFAVAVVKKSADL 834
```

Search completed: January 29, 2002, 10:24:08
Job time: 514 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:26:43 ; Search time 144.96 Seconds
(without alignments)
9.459 Million cell updates/sec

Title: US-09-763-397A-16
Perfect score: 100
Sequence: 1 DGNCEIPHVNEFSAILDL 18
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	622	2 A32499	apical membrane an
2	93	93.0	622	2 D44986	apical membrane an
3	93	93.0	622	2 B44986	apical membrane an
4	93	93.0	622	2 A44986	apical membrane an
5	93	93.0	622	2 C44986	apical membrane an
6	66	66.0	563	2 A39238	66K merozoite surf
7	63	63.0	562	2 A44944	apical membrane an
8	51	51.0	347	2 S67159	probable membrane
9	47.5	47.5	160	2 T01355	translation initia
10	45	45.0	183	2 A41351	multidrug resist
11	45	45.0	285	2 T27122	hypothetical prote
12	45	45.0	947	2 G70657	probable adi - Myc
13	45	45.0	1419	1 DVZOF	multidrug resist
14	44.5	44.5	967	2 T40171	probable chromosom
15	44	44.0	396	2 C84954	dihydropolipamide S
16	44	44.0	1804	2 T30563	resistance protein
17	44	44.0	1810	2 T30562	resistance protein
18	43.5	43.5	145	2 S21058	translation initia
19	43.5	43.5	158	2 F86272	initiation factor
20	43.5	43.5	159	2 S21060	translation initia
21	43.5	43.5	159	2 S21059	translation initia
22	43.5	43.5	160	2 T07133	translation initia
23	43.5	43.5	161	1 F7AAA	translation initia
24	43	43.0	827	2 C84546	probable transport
25	43	43.0	841	2 S69563	suppressor protein
26	43	43.0	1073	1 A25526	ring-infected eryt
27	42	42.0	238	2 T25955	hypothetical prote
28	42	42.0	312	2 S74669	hypothetical prote
29	42	42.0	922	2 T37256	metalloproteinase

ALIGNMENTS

RESULT 1

A32499 apical membrane antigen 1 - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 12-Oct-1989 #sequence_revision 12-Oct-1989 #text_change 09-Jun-2000
C:Accession: A32499
R:Peterson, M.G.; Marshall, V.M.; Smythe, J.A.; Crewther, P.E.; Lew, A.; Silva, A.; A Mol. Cell. Biol. 9, 3151-3154, 1989
A:Title: Integral membrane protein located in the apical complex of Plasmodium falcip
A:Reference number: A32499; MUID:89384584
A:Accession: A32499
A:Status: preliminary
A:Molecule type: DNA; mRNA
A:Residues: 1-622 <PEP>

A:Cross-references: GB:M27133; NID:gl60072; PID:gl60073
C:Keywords: membrane protein; surface antigen

Query Match 100.0%; Score 100; DB 2; Length 622;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DGNCEIPHVNEFSAILDL 18
|||||
Db 317 DGNCEIPHVNEFSAILDL 334

RESULT 2

D44986 apical membrane antigen 1 - malaria parasite (Plasmodium falciparum) (strain 7G8)
C:Species: Plasmodium falciparum
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000
C:Accession: D44986
R:Thomas, A.W.; Waters, A.P.; Carr, D.
Mol. Biochem. Parasitol. 42, 285-288, 1990
A:Title: Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate a
A:Reference number: A44986; MUID:91101665
A:Accession: D44986
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: DNA
A:Residues: 1-622 <THO>

A:Cross-references: GB:M34555
C:Keywords: membrane protein; surface antigen

Query Match 93.0%; Score 93; DB 2; Length 622;
Best Local Similarity 94.4%; Pred. No. 2.1e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DGNCEIPHVNEFSAILDL 18
|||||
Db 317 DGNCEIPHVNEFSAILDL 334

major surface glyc
protein kinase SSK
probable ubiquitin
probable membrane
conserved hypotet
hypothetical prote
probable flavoprot
glycoprotein B - h
ring-infected eryt
DNA (cytosine-5')-
probable retroelem
hypothetical prote
hypothetical prote
purine nucleotide
hypothetical prote
hypothetical prote

Mon Feb 4 15:23:38 2002

```

RESULT 3
B44986
apical membrane antigen 1 - malaria parasite (Plasmodium falciparum) (strain Thai Tn)
C:Species: Plasmodium falciparum
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000
C:Accession: B44986
R:Thomas, A.W.; Waters, A.P.; Carr, D.
Mol. Biochem. Parasitol. 42, 285-288, 1990
A:Title: Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate anti
A:Reference number: A44986; MUID:91101665
A:Accession: B44986
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: DNA
A:Residues: 1-622 <THO>
A:Cross-references: GB:M34553
C:Keywords: membrane protein; surface antigen

Query Match 93.0%; Score 93; DB 2; Length 622;
Best Local Similarity 94.4%; Pred. No. 2.le-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DGNCEIPHVNEFSAILD 18
DB 317 DGNCEIPHVNEFSANDL 334
|||||:|||||: ||

RESULT 4
A44986
apical membrane antigen 1 - malaria parasite (Plasmodium falciparum) (strain CAMP)
C:Species: Plasmodium falciparum
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000
C:Accession: A44986
R:Thomas, A.W.; Waters, A.P.; Carr, D.
Mol. Biochem. Parasitol. 42, 285-288, 1990
A:Title: Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate anti
A:Reference number: A44986; MUID:91101665
A:Accession: A44986
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: DNA
A:Residues: 1-622 <THO>
A:Cross-references: GB:M34552
C:Keywords: membrane protein; surface antigen

Query Match 93.0%; Score 93; DB 2; Length 622;
Best Local Similarity 94.4%; Pred. No. 2.le-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DGNCEIPHVNEFSAILD 18
DB 317 DGNCEIPHVNEFSANDL 334
|||||:|||||: ||

RESULT 5
C44986
apical membrane antigen 1 - malaria parasite (Plasmodium falciparum) (strain FCR 3)
C:Species: Plasmodium falciparum
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000
C:Accession: C44986
R:Thomas, A.W.; Waters, A.P.; Carr, D.
Mol. Biochem. Parasitol. 42, 285-288, 1990
A:Title: Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate anti
A:Reference number: A44986; MUID:91101665
A:Accession: C44986
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: DNA
A:Residues: 1-622 <THO>
A:Cross-references: GB:M34554
C:Keywords: membrane protein; surface antigen

Query Match 93.0%; Score 93; DB 2; Length 622;
Best Local Similarity 94.4%; Pred. No. 2.le-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DGNCEIPHVNEFSAILD 18
DB 317 DGNCEIPHVNEFSANDL 334
|||||:|||||: ||

RESULT 6
A39238
66k merozoite surface antigen precursor - Plasmodium knowlesi
C:Species: Plasmodium knowlesi
C:Date: 13-Sep-1991 #sequence_revision 13-Sep-1991 #text_change 07-Feb-1997
C:Accession: A39238
R:Waters, A.P.; Thomas, A.W.; Deans, J.A.; Mitchell, G.H.; Hudson, D.E.; Miller, L.H.
J. Biol. Chem. 265, 17974-17979, 1990
A:Title: A merozoite receptor protein from Plasmodium knowlesi is highly conserved an
A:Reference number: A39238; MUID:91009268
A:Accession: A39238
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-563 <WAT>
A:Cross-references: GB:J05631
C:Keywords: surface antigen

Query Match 66.0%; Score 66; DB 2; Length 563;
Best Local Similarity 66.7%; Pred. No. 0.0055;
Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 DGNCEIPHVNEFSAILD 18
DB 262 DGNCEIPYVNEVEARSL 279
|||||:|||||: |

RESULT 7
A44944
apical membrane antigen 1 precursor - Plasmodium fragile
C:Species: Plasmodium fragile
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-May-2000
C:Accession: A44944
R:Peterson, M.G.; Nguyen-Dinh, P.; Marshall, V.M.; Elliott, J.F.; Collins, W.E.; Ande
Mol. Biochem. Parasitol. 39, 279-284, 1990
A:Title: Apical membrane antigen of Plasmodium fragile.
A:Reference number: A44944; MUID:90205978
A:Accession: A44944
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-562 <PET>
A:Cross-references: GB:M29898

Query Match 63.0%; Score 63; DB 2; Length 562;
Best Local Similarity 61.1%; Pred. No. 0.017;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DGNCEIPHVNEFSAILD 18
DB 262 DGNCEIPYVDVQAKDL 279
|||||:|||||: |

RESULT 8
S67159
probable membrane protein YOR262w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein O5365
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 05-Nov-1999
C:Accession: S67159
R:Jauniaux, J.C.; Poiray, R.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67143
```


A:Accession: S67159
A:Molecule type: DNA
A:Residues: 1-347 <JAU>
A:Cross-references: EMBL:Z75170; NID:g1420590; PID:e252113; PID:g1420591; GSPDB:GN000015;
A:Experimental source: strain S288C
C:Genetics:
A:Gene: MIPS:YOR262W
A:Map position: 15R
C:Keywords: transmembrane protein
F:152-168/Domain: transmembrane #status predicted <TM>

Query Match 51.0%; Score 51; DB 2; Length 347;
Best Local Similarity 61.5%; Pred. No. 0.97;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 DIPHVNFEAIDL 18
|:|:|:|:|:|:|:
Db 167 DLPHINVFSLKIDM 179

RESULT 9
T01355
translation initiation factor eIF-5A [similarity] - maize
C:Species: Zea mays (maize)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 18-Aug-2000
C:Accession: T01355
R:Korn, A.M.; Baysdorfer, C.
submitted to the EMBL Data Library, November 1997
A:Reference number: Z14300
A:Accession: T01355
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-160 <KOR>
A:Cross-references: EMBL:AF034943; NID:g2668737; PIDN:AA88614.1; PID:g2668738
A:Experimental source: strain B73
C:Genetics:
A:Gene: TIF5A
C:Function:
A:Description: precise role in initiation of protein synthesis is not yet known; seems to
C:Superfamily: translation initiation factor eIF-5A
C:Keywords: hypusine; protein biosynthesis
F:53/Modified site: N6-(4-amino-2-hydroxybutyl)lysine (Lys) #status predicted

Query Match 47.5%; Score 47.5; DB 2; Length 160;
Best Local Similarity 61.1%; Pred. No. 1.6;
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 2;

QY 3 NCEDIPHN--EFSAILD 18
|| ||||| |:
Db 81 NC-DIPHNRTYQLIDI 97

RESULT 10
A41351
multidrug resistance protein mdr1 - malaria parasite (Plasmodium falciparum) (fragment)
C:Species: Plasmodium falciparum
C:Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 02-Feb-2001
C:Accession: A41351
R:Wilson, C.W.; Serrano, A.E.; Wasley, A.; Bogenschutz, M.P.; Shankar, A.H.; Wirth, D.F.
Science 244, 1184-1186, 1989
A:Title: Amplification of a gene related to mammalian mdr genes in drug-resistant Plasmo
A:Reference number: A41351; MUID:89266942
A:Accession: A41351
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-183 <WIL>
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
C:Keywords: ATP; duplication; glycoprotein; membrane protein; nucleotide binding; P-loop
F:1-183/Domain: ATP-binding cassette homology (fragment) <ABC2>
F:1-9/Region: nucleotide-binding motif A (P-loop)

F:172-176/Region: nucleotide-binding motif B

Query Match 45.0%; Score 45; DB 2; Length 183;
Best Local Similarity 47.1%; Pred. No. 4.7;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DGNCEIPHNFEAID 17
| | | | |:
Db 115 DATLEDVKRVSKFAAID 131

RESULT 11
T27122
hypothetical protein Y53C10A.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27122
R:White, S.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z20314
A:Accession: T27122
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-285 <WIL>
A:Cross-references: EMBL:AL033536; PIDN:CAA22143.1; CBSP:Y53C10A.7
A:Experimental source: clone Y53C10A
C:Genetics:
A:Gene: CBSP:Y53C10A.7
A:Introns: 12/2; 36/2; 61/3; 207/3

Query Match 45.0%; Score 45; DB 2; Length 285;
Best Local Similarity 45.0%; Pred. No. 7.7;
Matches 9; Conservative 5; Mismatches 2; Indels 4; Gaps 2;

QY 1 DGNCE--DIP--HVNFEFSAI 16
| | | | |:
Db 171 DGNCDFTDVPKHLSHFSRV 190

RESULT 12
G70657
probable adi - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: G70657
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987
A:Accession: G70657
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-947 <COL>
A:Cross-references: GB:Z83863; GB:AL123456; NID:g3261685; PIDN:CAB06177.1; PID:g29097
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: adi

Query Match 45.0%; Score 45; DB 2; Length 947;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 EDIPHNFEAIDL 18
|:|:|:|:|:|:|:
Db 789 QDLPHLPDFSEFDL 802

RESULT 13

DVZOF
multidrug resistance protein - malaria parasite (Plasmodium falciparum)
N:Alternate names: P-glycoprotein
C:Species: Plasmodium falciparum
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 19-Jan-2001
C:Accession: S18204; A32547
R:Triglia, T.; Foote, S.J.; Kemp, D.J.; Cowman, A.F.
Mol. Cell. Biol. 11, 5244-5250, 1991
A:Title: Amplification of the multidrug resistance gene pfmdr1 in Plasmodium falciparum
A:Reference number: S18204; MUID:92017800
A:Accession: S18204
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1419 <PRT>
A:Cross-references: EMBL:X56851; NID:99935; PIDN:CAA40180.1; PID:g9936
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1990
R:Footte, S.J.; Thompson, J.K.; Cowman, A.F.; Kemp, D.J.
Cell 57, 921-930, 1989
A:Title: Amplification of the multidrug resistance gene in some chloroquine-resistant is
A:Reference number: A32547; MUID:89288297
A:Accession: A32547
A:Molecule type: mRNA
A:Residues: 1-1419 <FOO>
A:Cross-references: GB:M29154; GB:M24322; NID:gl60398; PIDN:AAA29646.1; PID:gl60399
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembrane pr
F:1-632,655-1276/Region: duplication
F:1-632,655-1276/Region: duplication
F:396-638/Domain: ATP-binding cassette homology <ABC1>
F:413-420/Region: nucleotide-binding motif A (P-loop)
F:583-587/Region: nucleotide-binding motif B
F:1144-1387/Domain: ATP-binding cassette homology <ABC2>
F:1161-1168/Region: nucleotide-binding motif A (P-loop)
F:1332-1336/Region: nucleotide-binding motif B

Query Match 45.0%; Score 45; DB 1; Length 1419;

Best Local Similarity 47.1%; Pred. No. 45;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DGNCEIPHVNEFSAID 17

DB 1275 DATLEDVKRVKFAAID 1291

RESULT 14

T40171
probable chromosomal segregation protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T40171
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, August 1998
A:Reference number: Z21909
A:Accession: T40171
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-967 <LYN>
A:Cross-references: EMBL:AL031262; PIDN:CAA20318.1; GSPDB:GN00067; SPDB:SPBC30B4.05
A:Experimental source: strain 972h-; cosmid c30B4
C:Genetics:
A:Gene: SPDB:SPBC30B4.05
A:Map position: 2
A:Introns: 20/1

Query Match 44.5%; Score 44.5; DB 2; Length 967;

Best Local Similarity 50.0%; Pred. No. 36;
Matches 9; Conservative 3; Mismatches 1; Indels 5; Gaps 1;

QY 1 DGNCEIP-----HVNEF 13

DB 214 DLNCODIPEFFDHMSEF 231

RESULT 15

C84954
dihydrolipoamide S-acetyltransferase (EC 2.3.1.12) [imported] - Buchnera sp. (strain
C:Species: Buchnera sp.
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: C84954
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp
A:Reference number: A84930; MUID:20445173
A:Accession: C84954
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-396 <STO>
A:Cross-references: GB:AP000398; GSPDB:GN00144
A:Experimental source: strain APS
C:Genetics:
A:Gene: aceF; BU206
C:Keywords: acyltransferase; coenzyme A

Query Match 44.0%; Score 44; DB 2; Length 396;

Best Local Similarity 43.8%; Pred. No. 16;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 NCEDIPHVNEFSAIDL 18

DB 184 NWMNIPHTQFDEVDI 199

Search completed: January 29, 2002, 10:26:43
Job time: 654 sec

THIS PAGE BLANK (USPTO)

FT CONFLICT 503 N -> R (IN REF. 1: AAA29476).
SQ SEQUENCE 622 AA; 72009 MW; 7D41355E249FA18F CRC64;

Query Match 100.0%; Score 100; DB 1; Length 622;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNCEIDIPHVNEFSIDL 18
|||||
Db 317 DGNCEIDIPHVNEFSIDL 334

RESULT 2.
AMAL_PLAF8
ID AMAL_PLAF8 STANDARD; PRT; 622 AA.
AC P50492;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
GN AMA-1 OR PF83.
OS Plasmodium falciparum (isolate 768).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57266;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91101665; PubMed=2270110;
RA Thomas A.W., Waters A.P., Carr D.;
RT "Analysis of variation in pf83, an erythrocytic merozoite vaccine candidate antigen of Plasmodium falciparum";
RL Mol. Biochem. Parasitol. 42:285-287(1990).
CC -!- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO PK66 FROM P.KNOWLES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; M58545; AAA29718.1; -
DR InterPro: IPR003298; Apmem_Agl.
DR Pfam: PF02430; AMA-1; 1.
KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 622 APICAL MEMBRANE ANTIGEN 1.
FT DOMAIN 25 546 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 547 567 POTENTIAL.
FT DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 162 162 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 422 422 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 622 AA; 71943 MW; 26CE8CF6D07C637 CRC64;

Query Match 93.0%; Score 93; DB 1; Length 622;
Best Local Similarity 94.4%; Pred. No. 7.1e-08;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DGNCEIDIPHVNEFSIDL 18
|||||
Db 317 DGNCEIDIPHVNEFSIDL 334

RESULT 4
AMAL_PLAFG
ID AMAL_PLAFG STANDARD; PRT; 622 AA.
AC P50490;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
GN AMA-1 OR PF83.
OS Plasmodium falciparum (isolate FCR-3 / Gambia).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5838;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91101665; PubMed=2270110;

FT CONFLICT 503 N -> R (IN REF. 1: AAA29476).
SQ SEQUENCE 622 AA; 72009 MW; 7D41355E249FA18F CRC64;

Query Match 100.0%; Score 100; DB 1; Length 622;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNCEIDIPHVNEFSIDL 18
|||||
Db 317 DGNCEIDIPHVNEFSIDL 334

RESULT 2.
AMAL_PLAF8
ID AMAL_PLAF8 STANDARD; PRT; 622 AA.
AC P50492;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
GN AMA-1 OR PF83.
OS Plasmodium falciparum (isolate 768).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57266;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91101665; PubMed=2270110;
RA Thomas A.W., Waters A.P., Carr D.;
RT "Analysis of variation in pf83, an erythrocytic merozoite vaccine candidate antigen of Plasmodium falciparum";
RL Mol. Biochem. Parasitol. 42:285-287(1990).
CC -!- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO PK66 FROM P.KNOWLES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; M58548; AAA29721.1; -
DR InterPro: IPR003298; Apmem_Agl.
DR Pfam: PF02430; AMA-1; 1.
KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 622 APICAL MEMBRANE ANTIGEN 1.
FT DOMAIN 25 546 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 547 567 POTENTIAL.
FT DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 162 162 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 422 422 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 622 AA; 71990 MW; 1C9C8715D8E2915F CRC64;

Query Match 93.0%; Score 93; DB 1; Length 622;
Best Local Similarity 94.4%; Pred. No. 7.1e-08;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DGNCEIDIPHVNEFSIDL 18
|||||
Db 317 DGNCEIDIPHVNEFSIDL 334

RESULT 3

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL: M58547: AAA29720.1: "-
DR InterPro: IPR003298; Apmem_Ag1.
DR Pfam: PF02430; AMA-1; 1.
KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 622 APICAL MEMBRANE ANTIGEN 1.
FT DOMAIN 25 546 EXTRACELLULAR (POTENTIAL).
FT TRANSEM 547 567 POTENTIAL..
FT DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 422 422 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 622 AA; 71989 MW; 1FDFA53593C94CC5 CRC64;
Query Match 93.0%; Score 93; DB 1; Length 622;
Best Local Similarity 94.4%; Pred No. 7.1e-08;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 DGNCEDIPHVNEFSIDL 18
Db 317 DGNCEDIPHVNEFSANDL 334
|||||
RESULT 6
PK66_PLAKU ID PK66_PLAKU STANDARD; PRT; 563 AA.
AC P21303;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE MEROZOITE RECEPTOR PK66 PRECURSOR (66 KDA PROTECTIVE MINOR SURFACE
DE ANTIGEN).
GN PK66.
OS Plasmodium knowlesi (strain nuri).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5852;
RN [1]
RN RP SEQUENCE FROM N.A.
RN RC STRAIN=LINE W1.
RN RX MEDLINE=91009268; PubMed=2211675;
RA Waters A.P., Thomas A.W., Deans J.A., Mitchell G.H., Hudson D.E.,
RA Miller L.H., McCutchan T.F., Cohen S.;
RT "A merozoite receptor protein from Plasmodium knowlesi is highly
RT conserved and distributed throughout Plasmodium.";
RL J. Biol. Chem. 265:17974-17979(1990).
RN [2]
RN RP SEQUENCE FROM N.A.
RN RC STRAIN=NURI;
RN RX MEDLINE=91187058; PubMed=2011149;
RA Waters A.P., Thomas A.W., Mitchell G.H., McCutchan T.F.;
RA "Intra-generic conservation and limited inter-strain variation in a
RT protective minor surface antigen of Plasmodium knowlesi merozoites."
RL Mol. Biochem. Parasitol. 44:141-144(1991).
CC -1- FUNCTION: MEROZOITE RECEPTOR PK66 IS A SURFACE ANTIGEN INVOLVED
CC IN PARASITE INVASION OF ERYTHROCYTES.
CC CC SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CONCENTRATED AT THE
CC APICAL END PRIOR TO RUPTURE, FOLLOWING WHICH IT CAN DISTRIBUTE
CC ITSELF ENTIRELY ACROSS THE SURFACE OF THE FREE MEROZOITE. DURING
CC INVASION PK66 IS EXCLUDED FROM THE ERYTHROCYTE AT, AND BEHIND, THE
CC INVASION INTERFACE.
CC -1- MISCELLANEOUS: PK66 EXPRESSED IN MATURE SCHIZONTS IS RAPIDLY
CC PROCESSED AS THE SCHIZONT RUPTURES, YIELDING A 42/44 KDA DOUBLET
CC ASSOCIATED WITH THE SURFACE MEROZOITE.
CC -1- SIMILARITY: STRONG TO AMA-1/PF83 OF P.FALCIPARUM, P.CHABAUDI, AND
CC P.FRAGILE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; M58317; AAA63444.1; -;
 DR EMBL; M61097; AAA29728.1; -;
 DR InterPro: IPR003298; Apmem_Ag1.
 DR Pfam: PF02430; AMA-1; 1.
 KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
 FT SIGNAL 1 13 POTENTIAL.
 FT CHAIN 14 563 MERZOITE RECEPTOR PK66.
 FT DOMAIN 14 487 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 488 508 POTENTIAL.
 FT DOMAIN 509 563 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 107 107 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 228 228 N > K (IN STRAIN LINE W1).
 SQ SEQUENCE 563 AA; 64680 MW; 9EF0620264AA5CCF CRC64;

Query Match 66.0%; Score 66; DB 1; Length 563;

Best Local Similarity 66.7%; Pred. No. 0.002; Mismatches 2; Indels 0; Gaps 0;

Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 DGNCEIPHYNEFSIDL 18
 |||||:|||||
 DB 262 DGNCEIPYVNEARS 279

RESULT 7

ID AMA1_PLAFLR STANDARD; PRT; 562 AA.
 AC P22622;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MERZOITE SURFACE ANTIGEN).
 GN AMA-1 OR AG352.
 OS Plasmodium fragile.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5857;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90205978; PubMed=2181309;
 RA Peterson M.G., Nguyen-Dinh P., Marshall V.M., Elliott J.F.,
 RA Collins W.E., Anders R.F., Kemp D.J.;
 RT "Apical membrane antigen of Plasmodium fragile."
 RL Mol. Biochem. Parasitol. 39:279-284 (1990).
 CC -!- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FALCIPARUM, AND
 CC TO PK66 FROM P.KNOWLES.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; M29898; AAA29474.1; -;
 DR PIR; A44944; A44944.
 DR InterPro: IPR003298; Apmem_Ag1.
 DR Pfam: PF02430; AMA-1; 1.
 KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.

FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 562 APICAL MEMBRANE ANTIGEN 1.
 FT DOMAIN 22 484 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 485 507 POTENTIAL.
 FT DOMAIN 508 562 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 562 AA; 64488 MW; 9EAB72D437EA7164 CRC64;

Query Match 63.0%; Score 63; DB 1; Length 562;

Best Local Similarity 61.1%; Pred. No. 0.0063;

Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DGNCEIPHYNEFSIDL 18
 |||||:|||||
 DB 262 DGNCEIPYVQVQAKDL 279

RESULT 8

ID IF5A_MAIZE STANDARD; PRT; 160 AA.
 AC P80639; P56332;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE INITIATION FACTOR 5A (EIF-5A) (EIF-4D).
 GN TIF5A.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. B73;
 RA Korn A.M., Baysdorfer C.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. Al88;
 RA Cordts S., Loerz H., Dresselhaus T.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 71-78 AND 118-129.

RC TISSUE=Coleoptile;
 RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
 RA Pernollet J.-C., Zivy M., de Vienne D.;
 RT "The maize two dimensional gel protein database: towards an integrated
 RT genome analysis program."
 RL Theor. Appl. Genet. 93:997-1005 (1996).

CC -!- FUNCTION: THE PRECISE ROLE OF EIF-5A IN PROTEIN BIOSYNTHESIS IS
 CC NOT KNOWN BUT IT FUNCTIONS BY PROMOTING THE FORMATION OF THE FIRST
 CC PEPTIDE BOND.

CC -!- PFM: EIF-5A SEEMS TO BE THE ONLY EUKARYOTIC PROTEIN TO HAVE AN
 CC HYPSINE RESIDUE WHICH IS A POST-TRANSLATIONAL MODIFICATION OF A
 CC LYSINE BY THE ADDITION OF A BUTYLAMINO GROUP (FROM SPERMIDINE).

CC -!- SIMILARITY: BELONGS TO THE EIF-5A FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AF034943; AAB88614.1; -;
 DR EMBL; Y07920; CAA69225.1; -;
 DR Maize-2DPAGE; P80639; COLEOPTILE.

DR MaizeDB; 123964; --
DR InterPro: IPR001884; EIF5A_hypusine.
DR Pfam: PF01287; eIF-5a; 1.
DR PROSITE; PS00302; IF5A_HYPUSINE; 1.
KW Protein biosynthesis; Initiation factor; Hypusine.
FT MOD_RES 53 53 HYPUSINE (BY SIMILARITY).
SQ SEQUENCE 160 AA; 17497 MW; 7CCFADF2EC09CAD CRC64;

Query Match 47.5%; Score 47.5; DB 1; Length 160;
Best Local Similarity 61.1%; Pred. No. 0.66; 2; Indels 3; Gaps 2;
Matches 11; Conservative 2; Mismatches 2;

Oy 3 NCEDIPHYN--EFSALD 18
Db 81 NC-DIPHNRTEYQLIDI 97
||| ||||| | : ||:
||| ||||| | : ||:

RESULT 9
MDR_PLAFF STANDARD; PRT; 1419 AA.
AC P13568;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE MULTIDRUG RESISTANCE PROTEIN (CHLOROQUINE RESISTANCE PROTEIN).
GN MDR1.
OS Plasmodium falciparum (isolate FC27 / Papua New Guinea).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5837;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89288297; PubMed=2701941;
RA Foote S.J., Thompson J.K., Cowman A.F., Kemp D.J.;
RT "Amplification of the multidrug resistance gene in some chloroquine-
resistant isolates of P. falciparum";
RL Cell 57:921-930(1989).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92017800; PubMed=1922044;
RA Triglia T., Foote S.J., Kemp D.J., Cowman A.F.;
RT "Amplification of the multidrug resistance gene pfmdr1 in Plasmodium
falciparum has arisen as multiple independent events";
RL Mol. Cell. Biol. 11:5244-5250(1991).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93149200; PubMed=84266508;
RA Wilson C.M., Volkman S.K., Thaithong S., Martin R.K., Kyle D.E.,
RA Milhous W.K., Wirth D.F.;
RT "Amplification of pfmdr1 associated with mefloquine and halofantrine
resistance in Plasmodium falciparum from Thailand";
RL Mol. Biochem. Parasitol. 57:151-160(1993).
CC -!- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- MISCELLANEOUS: P. FALCIPARUM RESISTANT TO THE DRUG CHLOROQUINE.
CC HAVE MULTIPLE COPIES OF THE GENE CODING FOR MDR.
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MDR SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M29154; AAA29646.1; -;
DR EMBL; X56851; CAA40180.1; -;
DR EMBL; S53996; AAD13870.1; -;
DR PIR; A32547; DVZQF.
DR PIR; S18204; S18204.

DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001140; ABC_transportr_tmem.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR SMART; SMO0382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat.
FT DOMAIN 1 55 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 56 82 POTENTIAL.
FT TRANSMEM 91 116 POTENTIAL.
FT TRANSMEM 160 188 POTENTIAL.
FT TRANSMEM 194 212 POTENTIAL.
FT TRANSMEM 279 298 POTENTIAL.
FT TRANSMEM 314 338 POTENTIAL.
FT DOMAIN 339 788 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 789 807 POTENTIAL.
FT TRANSMEM 825 846 POTENTIAL.
FT TRANSMEM 908 928 POTENTIAL.
FT TRANSMEM 1028 1048 POTENTIAL.
FT TRANSMEM 1063 1083 POTENTIAL.
FT DOMAIN 1084 1419 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 643 661 POLY-ASN.
FT NP_BIND 413 420 ATP (POTENTIAL).
FT NP_BIND 1161 1168 ATP (POTENTIAL).
FT REPEAT 1 721
FT REPEAT 722 1419
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 258 258 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 964 964 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1419 AA; 162251 MW; 0F96C7C1850B33D0 CRC64;

Query Match 45.0%; Score 45; DB 1; Length 1419;
Best Local Similarity 47.1%; Pred. No. 17;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Oy 1 DGNCEIDPHVNEFSALD 17
Db 1275 DATLEDVRVSKFAALD 1291
||| :||:|:|:
||| :||:|:|:

RESULT 10
ODP2_BUCAI STANDARD; PRT; 396 AA.
AC P57302;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DIHYDROLIPOAMIDE ACETYLTRANSFERASE COMPONENT OF PYRUVATE DEHYDROGENASE
DE COMPLEX (EC 2.3.1.12) (E2).
GN ACEF OR BUZ06.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=TORYO 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS";
RL Nature 407:81-86(2000).
CC -!- FUNCTION: THE PYRUVATE DEHYDROGENASE COMPLEX CATALYZES THE OVERALL
CC CONVERSION OF PYRUVATE TO ACETYL-COA & CO(2). IT CONTAINS MULTIPLE
CC COPIES OF THREE ENZYMATIC COMPONENTS: PYRUVATE DEHYDROGENASE (E1),
CC DIHYDROLIPOAMIDE ACETYLTRANSFERASE (E2) & LIPOAMIDE DEHYDROGENASE
CC (E3) (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ACETYL-COA + DIHYDROLIPOAMIDE -> COA +
CC S-ACETYLDIHYDROLIPOAMIDE.
CC -!- COFACTOR: THE E2 COMPONENT CONTAINS ONE COVALENTLY-BOUND LIPOYL

CC COFACTOR (BY SIMILARITY).
CC -!- SUBUNIT: FORMS A 24-POLYPEPTIDE STRUCTURAL CORE WITH OCTAHEDRAL
CC SYMMETRY (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.
CC -!- SIMILARITY: CONTAINS 1 LIPOYL-BINDING DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP001118; BAB12923.1; -.
CC InterPro: IPR001078; 2oxoacid_dh.
CC InterPro: IPR000089; Biotin_lipoyl.
CC InterPro: IPR003016; Lipoyl.
CC Pfam: PF00198; 2-oxoacid_dh; 1.
CC ProDom: PD001115; 2oxoacid_dh; 1.
CC PROSITE; PS00189; LIPOYL; 1.
CC Glycolysis; Transferase; Acyltransferase; Lipoyl; Complete proteome.
KW BINDING 35 35 LIPOYL (BY SIMILARITY).
FT ACT_SITE 369 369 POTENTIAL.
SQ SEQUENCE 396 AA; 45277 MW; 1D5B49322AE963BC CRC64;

Query Match 44.0%; Score 44; DB 1; Length 396;
Best Local Similarity 43.8%; Pred. No. 6.4;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 NCEDIPHVNFSAIDL 18
| :||| : : :
Db 184 NWNIPHTQFDEVDI 199

RESULT 11
IF51_NICPL
ID IF51_NICPL STANDARD; PRT; 159 AA.
AC P24921;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE INITIATION FACTOR 5A-1 (EIF-5A) (EIF-4D).
GN EIF-5A1.
OS Nicotiana plumbaginifolia (Leadwort-leaved tobacco), and
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4092, 4097;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=92178957; PubMed=1542563;
RA Chamot D., Kuhlmeier C.;
RT "Differential expression of genes encoding the hypusine-containing
translation initiation factor, eIF-5A, in tobacco.";
RL Nucleic Acids Res. 20:665-669(1992).
CC -!- FUNCTION: THE PRECISE ROLE OF EIF-5A IN PROTEIN BIOSYNTHESIS IS
CC NOT KNOWN BUT IT FUNCTIONS BY PROMOTING THE FORMATION OF THE FIRST
CC PEPTIDE BOND.
CC -!- PTM: EIF-5A SEEMS TO BE THE ONLY EUKARYOTIC PROTEIN TO HAVE AN
CC HYPUSINE RESIDUE WHICH IS A POST-TRANSLATIONAL MODIFICATION OF A
CC LYSINE BY THE ADDITION OF A BUTYLAMINO GROUP (FROM SPERMIDINE).
CC -!- MISCELLANEOUS: THERE ARE AT LEAST TWO GENES FOR EIF-5A IN TOBACCO:
CC 5A1 MAY REGULATE THE LIGHT-DEPENDENT TRANSLATION OF SPECIFIC
CC TRANSCRIPTS WHILE 5A2 MAY BE A HOUSEKEEPING PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE EIF-5A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X63542; CAA45104.1; -.
CC InterPro: IPR001884; EIF5A_hypusine.
CC Pfam: PF01287; eIF-5a; 1.
CC PROSITE; PS00302; IF5A_HYPUSINE; 1.
CC -----

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X63543; CAA45105.1; -.
CC EMBL; X63541; CAA45103.1; -.
CC PIR; S21060; S21060.
CC PIR; S21058; S21058.
CC InterPro: IPR001884; EIF5A_hypusine.
CC Pfam: PF01287; eIF-5a; 1.
CC PROSITE; PS00302; IF5A_HYPUSINE; 1.
KW Protein biosynthesis; Initiation factor; Hypusine; Multigene family.
FT MOD_RES 52 52 HYPUSINE (BY SIMILARITY).
SQ SEQUENCE 159 AA; 17370 MW; 5174BC5809BE275 CRC64;

Query Match 43.5%; Score 43.5; DB 1; Length 159;
Best Local Similarity 50.0%; Pred. No. 3;
Matches 9; Conservative 4; Mismatches 2; Indels 3; Gaps 2;

QY 3 NCEDIPHVNFSAIDL 18
| :||| : : :
Db 80 NC-DVPHVNRDTYQLIDI 96

RESULT 12
IF52_NICPL
ID IF52_NICPL STANDARD; PRT; 159 AA.
AC P24922;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE INITIATION FACTOR 5A-2 (EIF-5A) (EIF-4D).
GN EIF-5A2.
OS Nicotiana plumbaginifolia (Leadwort-leaved tobacco).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4092;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=92178957; PubMed=1542563;
RA Chamot D., Kuhlmeier C.;
RT "Differential expression of genes encoding the hypusine-containing
translation initiation factor, eIF-5A, in tobacco.";
RL Nucleic Acids Res. 20:665-669(1992).
CC -!- FUNCTION: THE PRECISE ROLE OF EIF-5A IN PROTEIN BIOSYNTHESIS IS
CC NOT KNOWN BUT IT FUNCTIONS BY PROMOTING THE FORMATION OF THE FIRST
CC PEPTIDE BOND.
CC -!- PTM: EIF-5A SEEMS TO BE THE ONLY EUKARYOTIC PROTEIN TO HAVE AN
CC HYPUSINE RESIDUE WHICH IS A POST-TRANSLATIONAL MODIFICATION OF A
CC LYSINE BY THE ADDITION OF A BUTYLAMINO GROUP (FROM SPERMIDINE).
CC -!- MISCELLANEOUS: THERE ARE AT LEAST TWO GENES FOR EIF-5A IN TOBACCO:
CC 5A1 MAY REGULATE THE LIGHT-DEPENDENT TRANSLATION OF SPECIFIC
CC TRANSCRIPTS WHILE 5A2 MAY BE A HOUSEKEEPING PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE EIF-5A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X63542; CAA45104.1; -.
CC PIR; S21059; S21059.
CC InterPro: IPR001884; EIF5A_hypusine.
CC Pfam: PF01287; eIF-5a; 1.
CC PROSITE; PS00302; IF5A_HYPUSINE; 1.
CC -----

```

KW Protein biosynthesis; Initiation factor; Hypusine; Multigene family.
FT MOD_RES 52 HYPUSINE (BY SIMILARITY).
SQ SEQUENCE 159 AA; 17363 MW; 69B7BF6322429D91 CRC64;

Query Match 43.5%; Score 43.5; DB 1; Length 159;
Best Local Similarity 50.0%; Pred. No. 3;
Matches 9; Conservative 4; Mismatches 2; Indels 3; Gaps 2;

Qy 3 NCEDIPHN--EFSAILD 18
|| 1:|||| :|:
Db 80 NC-DVPHVNRDYLQIDI 96

RESULT 13
IF54_SOLTU STANDARD; PRT; 159 AA.
AC P56336; 1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE INITIATION FACTOR 5A-4 (EIF-5A) (EIF-4D).
GN EIF5A4.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. IRISH COBBLER;
RA In J.G., Fujino K., Kikuta Y.;
RT "Nucleotide sequence of five cDNAs encoding eukaryotic translation
  initiation factor 5A (eIF-5A) from potato.";
RL (In) Plant Gene Register PCR97-147.
CC -1- FUNCTION: THE PRECISE ROLE OF EIF-5A IN PROTEIN BIOSYNTHESIS IS
CC NOT KNOWN BUT IT FUNCTIONS BY PROMOTING THE FORMATION OF THE FIRST
CC PEPTIDE BOND.
CC -1- PTM: EIF-5A SEEMS TO BE THE ONLY EUKARYOTIC PROTEIN TO HAVE AN
CC HYPUSINE RESIDUE WHICH IS A POST-TRANSLATIONAL MODIFICATION OF A
CC LYSINE BY THE ADDITION OF A BUTYLAMINO GROUP (FROM SPERMIDINE).
CC -1- SIMILARITY: BELONGS TO THE EIF-5A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB004825; BAA20878.1;
CC InterPro: IPR001884; EIF5A_hypusine.
CC Pfam: PF01287; eif-5a; 1.
CC PROSITE: PS00302; IF5A_HYPUSINE; 1.
CC Protein biosynthesis; Initiation factor; Hypusine; Multigene family.
FT MOD_RES 52 HYPUSINE (BY SIMILARITY).
SQ SEQUENCE 159 AA; 17318 MW; D3DFD223969E18E9 CRC64;

Query Match 43.5%; Score 43.5; DB 1; Length 159;
Best Local Similarity 50.0%; Pred. No. 3;
Matches 9; Conservative 4; Mismatches 2; Indels 3; Gaps 2;

Qy 3 NCEDIPHN--EFSAILD 18
|| 1:|||| :|:
Db 80 NC-DVPHVNRDYLQIDI 96

RESULT 14
IF55_SOLTU STANDARD; PRT; 159 AA.
AC P56337;

```

```

DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE INITIATION FACTOR 5A-5 (EIF-5A) (EIF-4D).
GN EIF5A5.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. IRISH COBBLER;
RA In J.G., Fujino K., Kikuta Y.;
RT "Nucleotide sequence of five cDNAs encoding eukaryotic translation
  initiation factor 5A (eIF-5A) from potato.";
RL (In) Plant Gene Register PCR97-147.
CC -1- FUNCTION: THE PRECISE ROLE OF EIF-5A IN PROTEIN BIOSYNTHESIS IS
CC NOT KNOWN BUT IT FUNCTIONS BY PROMOTING THE FORMATION OF THE FIRST
CC PEPTIDE BOND.
CC -1- PTM: EIF-5A SEEMS TO BE THE ONLY EUKARYOTIC PROTEIN TO HAVE AN
CC HYPUSINE RESIDUE WHICH IS A POST-TRANSLATIONAL MODIFICATION OF A
CC LYSINE BY THE ADDITION OF A BUTYLAMINO GROUP (FROM SPERMIDINE).
CC -1- SIMILARITY: BELONGS TO THE EIF-5A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB004826; BAA20879.1;
CC InterPro: IPR001884; EIF5A_hypusine.
CC Pfam: PF01287; eif-5a; 1.
CC PROSITE: PS00302; IF5A_HYPUSINE; 1.
CC Protein biosynthesis; Initiation factor; Hypusine; Multigene family.
FT MOD_RES 52 HYPUSINE (BY SIMILARITY).
SQ SEQUENCE 159 AA; 17345 MW; 59D7B0C599FC5D2F CRC64;

Query Match 43.5%; Score 43.5; DB 1; Length 159;
Best Local Similarity 50.0%; Pred. No. 3;
Matches 9; Conservative 4; Mismatches 2; Indels 3; Gaps 2;

Qy 3 NCEDIPHN--EFSAILD 18
|| 1:|||| :|:
Db 80 NC-DVPHVNRDYLQIDI 96

RESULT 15
IF51_SOLTU STANDARD; PRT; 160 AA.
AC P56333; P56334;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE INITIATION FACTOR 5A (EIF-5A) (EIF-4D).
GN EIF5A1 AND EIF5A2.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. IRISH COBBLER;
RA In J.G., Fujino K., Kikuta Y.;
RT "Nucleotide sequence of five cDNAs encoding eukaryotic translation
  initiation factor 5A (eIF-5A) from potato.";
RL (In) Plant Gene Register PCR97-147.
CC -1- FUNCTION: THE PRECISE ROLE OF EIF-5A IN PROTEIN BIOSYNTHESIS IS

```

CC NOT KNOWN BUT IT FUNCTIONS BY PROMOTING THE FORMATION OF THE FIRST
CC PEPTIDE BOND.
CC -1- PTM: EIF-5A SEEMS TO BE THE ONLY EUKARYOTIC PROTEIN TO HAVE AN
CC HYPUSINE RESIDUE WHICH IS A POST-TRANSLATIONAL MODIFICATION OF A
CC LYSINE BY THE ADDITION OF A BUTYLAMINO GROUP (FROM SPERMIDINE).
CC -1- SIMILARITY: BELONGS TO THE EIF-5A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB004827; BAA20880.1; -;
CC EMBL; AB004823; BAA20876.1; -;
CC InterPro: IPR001884; EIF5A_hypusine.
CC Pfam: PF01287; eIF-5a; 1.
CC PROSITE; PS00302; IF5A_HYPUSINE; 1.
CC KW Protein biosynthesis; Initiation factor; Hypusine; Multigene family.
CC MOD_RES 52 52 HYPUSINE (BY SIMILARITY).
CC SQ SEQUENCE 160 AA; 17431 MW; E35BDF67185B9551 CRC64;
CC -----
Query Match 43.5%; Score 43.5; DB 1; Length 160;
Best Local Similarity 50.0%; Pred. No. 3;
Matches 9; Conservative 4; Mismatches 2; Indels 3; Gaps 2;
QY 3 NCEDIPHVN--EFSADL 18
DB 80 NC-DVPHVNRDYLQIDI 96
|||:|||||::|||:

Search completed: January 29, 2002, 11:13:45
Job time: 821 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:12:13 ; Search time 285.36 Seconds
(without alignments)
9.227 Million cell updates/sec

Title: us-09-763-397A-16

Perfect score: 100

Sequence: 1 DGNCEIPHVNEFSAIDL 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	100	100.0	437	5 Q9BHR3	Q9bhr3 plasmodium
2	100	100.0	437	5 Q9BHR2	Q9bhr2 plasmodium
3	100	100.0	437	5 Q9BHP6	Q9bhp6 plasmodium
4	100	100.0	437	5 Q9BHP2	Q9bhp2 plasmodium
5	100	100.0	437	5 Q9BH26	Q9bh26 plasmodium
6	100	100.0	526	5 Q9N9G0	Q9n9g0 plasmodium
7	100	100.0	526	5 Q9N9F2	Q9n9f2 plasmodium
8	100	100.0	526	5 Q9N9E4	Q9n9e4 plasmodium
9	100	100.0	526	5 Q9N9E1	Q9n9e1 plasmodium
10	100	100.0	526	5 Q9N9E0	Q9n9e0 plasmodium
11	95	95.0	437	5 Q9BHS1	Q9bhs1 plasmodium
12	95	95.0	437	5 Q9BHR5	Q9bhr5 plasmodium
13	95	95.0	437	5 Q9BHR1	Q9bhr1 plasmodium
14	95	95.0	437	5 Q9BHQ9	Q9bhq9 plasmodium
15	95	95.0	437	5 Q9BH08	Q9bhq8 plasmodium
16	95	95.0	437	5 Q9BHP7	Q9bhp7 plasmodium
17	95	95.0	437	5 Q9BHN8	Q9bhn8 plasmodium
18	95	95.0	437	5 Q9BHB6	Q9bhb6 plasmodium
19	95	95.0	437	5 Q9BH63	Q9bh63 plasmodium

20	95	95.0	526	5 Q9N9E8	Q9n9e8 plasmodium
21	95	95.0	526	5 Q9N9E6	Q9n9e6 plasmodium
22	95	95.0	622	5 Q94661	Q94661 plasmodium
23	93	93.0	402	5 Q9BIM8	Q9bim8 plasmodium
24	93	93.0	437	5 Q9BHS0	Q9bhs0 plasmodium
25	93	93.0	437	5 Q9BHR9	Q9bhr9 plasmodium
26	93	93.0	437	5 Q9BHR8	Q9bhr8 plasmodium
27	93	93.0	437	5 Q9BHR7	Q9bhr7 plasmodium
28	93	93.0	437	5 Q9BHR6	Q9bhr6 plasmodium
29	93	93.0	437	5 Q9BHR4	Q9bhr4 plasmodium
30	93	93.0	437	5 Q9BHR0	Q9bhr0 plasmodium
31	93	93.0	437	5 Q9BHQ7	Q9bhq7 plasmodium
32	93	93.0	437	5 Q9BHQ6	Q9bhq6 plasmodium
33	93	93.0	437	5 Q9BHQ5	Q9bhq5 plasmodium
34	93	93.0	437	5 Q9BHQ4	Q9bhq4 plasmodium
35	93	93.0	437	5 Q9BHQ3	Q9bhq3 plasmodium
36	93	93.0	437	5 Q9BHQ2	Q9bhq2 plasmodium
37	93	93.0	437	5 Q9BHQ0	Q9bhq0 plasmodium
38	93	93.0	437	5 Q9BHP9	Q9bhp9 plasmodium
39	93	93.0	437	5 Q9BHP8	Q9bhp8 plasmodium
40	93	93.0	437	5 Q9BHP5	Q9bhp5 plasmodium
41	93	93.0	437	5 Q9BHP4	Q9bhp4 plasmodium
42	93	93.0	437	5 Q9BHP3	Q9bhp3 plasmodium
43	93	93.0	437	5 Q9BHP0	Q9bhp0 plasmodium
44	93	93.0	437	5 Q9BHN9	Q9bhn9 plasmodium
45	93	93.0	437	5 Q9BHN7	Q9bhn7 plasmodium

ALIGNMENTS

RESULT 1

Q9BHR3 PRELIMINARY; PRT: 437 AA.
AC Q9BHR3:
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIGERIAN 036:
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ408311; CAC34752.1; -.
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 50220 MW; 1CBDF1C89F0C06B0 CRC64;

Query Match 100.0%; Score 100; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 9e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNCEIPHVNEFSAIDL 18
Db 171 DGNCEIPHVNEFSAIDL 188
|||||

RESULT 2

Q9BHR2 PRELIMINARY; PRT: 437 AA.
ID Q9BHR2:
AC Q9BHR2:
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).

```
GN AMA1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIGERIAN 039;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ408312; CAC34753.1;
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 50356 MW; 734A74080DIBAC7F CRC64;

Query Match 100.0%; Score 100; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 9e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNCEIPHVNEFSIDL 18
Db 171 DGNCEIPHVNEFSIDL 188

RESULT 3
ID Q9BHP6 PRELIMINARY; PRT; 437 AA.
AC Q9BHP6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIGERIAN 088;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ408334; CAC34775.1;
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 50288 MW; F0756EE4695FFAC CRC64;

Query Match 100.0%; Score 100; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 9e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNCEIPHVNEFSIDL 18
Db 171 DGNCEIPHVNEFSIDL 188

RESULT 4
ID Q9BHP2 PRELIMINARY; PRT; 437 AA.
AC Q9BHP2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=NIGERIAN 100;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ408340; CAC34781.1;
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 50311 MW; 57D2DFEB99235EFA CRC64;

Query Match 100.0%; Score 100; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 9e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNCEIPHVNEFSIDL 18
Db 171 DGNCEIPHVNEFSIDL 188

RESULT 5
ID Q9BH26 PRELIMINARY; PRT; 437 AA.
AC Q9BH26;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIGERIAN 016, AND NIGERIAN 096;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ408304; CAC34745.1;
DR EMBL; AJ408339; CAC34780.1;
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 50240 MW; F35E75801A46F3BD CRC64;

Query Match 100.0%; Score 100; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 9e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNCEIPHVNEFSIDL 18
Db 171 DGNCEIPHVNEFSIDL 188

RESULT 6
ID Q9N9G0 PRELIMINARY; PRT; 526 AA.
AC Q9N9G0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN 83/AMA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C9-6.1;
RX MEDLINE=20416492; PubMed=10960173;
RA Kocken C.H.M., Narum D.L., Massougboodji A., Ayivi B., Dubbeld M.A.,
```

```

RA Van der Wel A., Conway D.J., Sanni A., Thomas A.W.;
RT "Molecular characterisation of Plasmodium reichenowi apical membrane
RT antigen-1 (AMA-1), comparison with P. falciparum AMA-1, and antibody-
RT mediated inhibition of red cell invasion.";
RL Mol. Biochem. Parasitol. 109:147-156(2000).
DR EMBL: AJ271169; CAB97181.1; -.
DR InterPro: IPR003298; Apmem_Ag1.
DR Pfam: PF02430; AMA-1; 1.
DR PRINTS: PR01361; MEROZOITESA.
FT NON_TER 1
FT NON_TER 526
SQ SEQUENCE 526 AA; 60798 MW; 5654B6461BACF5ED CRC64;

Query Match 100.0%; Score 100; DB 5; Length 526;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DGNCEIDPHVNEFSAIDL 18
Db 309 DGNCEIDPHVNEFSAIDL 326

RESULT 7
Q9N9F2 ID Q9N9F2 PRELIMINARY; PRT; 526 AA.
AC Q9N9F2
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN 83/AMA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12-4.1;
RX MEDLINE=20416492; PubMed=10960173;
RA Kocken C.H.M., Narum D.L., Massougbodji A., Ayivi B., Dubbeld M.A.,
RA van der Wel A., Conway D.J., Sanni A., Thomas A.W.;
RT "Molecular characterisation of Plasmodium reichenowi apical membrane
RT antigen-1 (AMA-1), comparison with P. falciparum AMA-1, and antibody-
RT mediated inhibition of red cell invasion.";
RL Mol. Biochem. Parasitol. 109:147-156(2000).
DR EMBL: AJ271177; CAB97189.1; -.
DR InterPro: IPR003298; Apmem_Ag1.
DR Pfam: PF02430; AMA-1; 1.
DR PRINTS: PR01361; MEROZOITESA.
FT NON_TER 1
FT NON_TER 526
SQ SEQUENCE 526 AA; 60738 MW; 4A3F6DAA0BFA2573 CRC64;

Query Match 100.0%; Score 100; DB 5; Length 526;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DGNCEIDPHVNEFSAIDL 18
Db 309 DGNCEIDPHVNEFSAIDL 326

RESULT 8
Q9N9E4 ID Q9N9E4 PRELIMINARY; PRT; 526 AA.
AC Q9N9E4
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN 83/AMA-1.
OS Plasmodium falciparum.

```

```

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7-9.1;
RX MEDLINE=20416492; PubMed=10960173;
RA Kocken C.H.M., Narum D.L., Massougbodji A., Ayivi B., Dubbeld M.A.,
RA van der Wel A., Conway D.J., Sanni A., Thomas A.W.;
RT "Molecular characterisation of Plasmodium reichenowi apical membrane
RT antigen-1 (AMA-1), comparison with P. falciparum AMA-1, and antibody-
RT mediated inhibition of red cell invasion.";
RL Mol. Biochem. Parasitol. 109:147-156(2000).
DR EMBL: AJ271185; CAB97197.1; -.
DR InterPro: IPR003298; Apmem_Ag1.
DR Pfam: PF02430; AMA-1; 1.
DR PRINTS: PR01361; MEROZOITESA.
FT NON_TER 1
FT NON_TER 526
SQ SEQUENCE 526 AA; 60917 MW; C6522F8B072078D2D CRC64;

```

```

Query Match 100.0%; Score 100; DB 5; Length 526;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 DGNCEIDPHVNEFSAIDL 18
Db 309 DGNCEIDPHVNEFSAIDL 326

```

```

RESULT 9
Q9N9E1 ID Q9N9E1 PRELIMINARY; PRT; 526 AA.
AC Q9N9E1
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN 83/AMA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=27-2.1;
RX MEDLINE=20416492; PubMed=10960173;
RA Kocken C.H.M., Narum D.L., Massougbodji A., Ayivi B., Dubbeld M.A.,
RA van der Wel A., Conway D.J., Sanni A., Thomas A.W.;
RT "Molecular characterisation of Plasmodium reichenowi apical membrane
RT antigen-1 (AMA-1), comparison with P. falciparum AMA-1, and antibody-
RT mediated inhibition of red cell invasion.";
RL Mol. Biochem. Parasitol. 109:147-156(2000).
DR EMBL: AJ271188; CAB97200.1; -.
DR InterPro: IPR003298; Apmem_Ag1.
DR Pfam: PF02430; AMA-1; 1.
DR PRINTS: PR01361; MEROZOITESA.
FT NON_TER 1
FT NON_TER 526
SQ SEQUENCE 526 AA; 60730 MW; 997A82E10E85EF6C CRC64;

```

```

Query Match 100.0%; Score 100; DB 5; Length 526;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 DGNCEIDPHVNEFSAIDL 18
Db 309 DGNCEIDPHVNEFSAIDL 326

```

```

RESULT 10
Q9N9E0 ID Q9N9E0 PRELIMINARY; PRT; 526 AA.

```

```
AC Q9N9E0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN 83/AMA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=25-6.1;
RX MEDLINE=20416492; PubMed=10960173;
RA Kocken C.H.M., Narum D.L., Massouh-Bodji A., Ayivi B., Dubbeld M.A.,
RA van der Wel A., Conway D.J., Sanni A., Thomas A.W.;
RT "Molecular characterization of Plasmodium reichenowi apical membrane
RT antigen-1 (AMA-1), comparison with P. falciparum AMA-1, and antibody-
RT mediated inhibition of red cell invasion.";
RL Mol. Biochem. Parasitol. 109:147-156(2000).
DR EMBL; AJ271189; CAB97201.1; -
DR InterPro; IPR003298; Apmem_Ag1.
DR Pfam; PF02430; AMA-1; 1.
DR PRINTS; PRO1361; MEROZOITESA.
FT NON_TER 1
FT NON_TER 526
FT SEQUENCE 526 AA; 60727 MW; 934A9EFD7B93DF91 CRC64;

Query Match 100.0%; Score 100; DB 5; Length 526;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNCEIPHVNEFSIDL 18
DB 309 DGNCEIPHVNEFSIDL 326

RESULT 11
Q9BHS1
ID Q9BHS1 PRELIMINARY; PRT; 437 AA.
AC Q9BHS1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIGERIAN 002;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical
RT Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ408300; CAC34741.1; -
FT NON_TER 1
FT NON_TER 437
FT SEQUENCE 437 AA; 50218 MW; C043561972E8D49F CRC64;

Query Match 95.0%; Score 95; DB 5; Length 437;
Best Local Similarity 94.4%; Pred. No. 6.3e-08;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DGNCEIPHVNEFSIDL 18
DB 171 DGNCEIPHVNEFSIDL 188

RESULT 12
Q9BHS5
ID Q9BHS5 PRELIMINARY; PRT; 437 AA.
AC Q9BHS5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIGERIAN 034B;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical
RT Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ408309; CAC34750.1; -
FT NON_TER 1
FT NON_TER 437
FT SEQUENCE 437 AA; 50206 MW; EB05396E7D92AC98 CRC64;

Query Match 95.0%; Score 95; DB 5; Length 437;
Best Local Similarity 94.4%; Pred. No. 6.3e-08;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DGNCEIPHVNEFSIDL 18
DB 171 DGNCEIPHVNEFSIDL 188

RESULT 13
Q9BHR1
ID Q9BHR1 PRELIMINARY; PRT; 437 AA.
AC Q9BHR1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIGERIAN 044;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical
RT Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ408314; CAC34755.1; -
FT NON_TER 1
FT NON_TER 437
FT SEQUENCE 437 AA; 50232 MW; BA3061373049C9CE CRC64;

Query Match 95.0%; Score 95; DB 5; Length 437;
Best Local Similarity 94.4%; Pred. No. 6.3e-08;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DGNCEIPHVNEFSIDL 18
DB 171 DGNCEIPHVNEFSIDL 188

RESULT 14
Q9BHQ9
ID Q9BHQ9 PRELIMINARY; PRT; 437 AA.
AC Q9BHQ9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
```


DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
 GN AMA1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIGERIAN 050;
 RA Polley S.D., Conway D.J.;
 RT "Diversifying selection in domains of Plasmodium falciparum Apical
 Membrane Antigen 1 (AMA1).";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ408316; CAC34757.1; -.
 FT NON_TER 1
 FT NON_TER 437 437
 SQ SEQUENCE 437 AA; 50140 MW; 9FE785C026BEDC63 CRC64;

Query Match 95.0%; Score 95; DB 5; Length 437;
 Best Local Similarity 94.4%; Pred. No. 6.3e-08;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DGNCEIDPHVNEFSAIDL 18
 |||
 Db 171 DGNCEIDPHVNEFPAIDL 188

RESULT 15
 Q9BHQ8
 ID Q9BHQ8 PRELIMINARY; PRT; 437 AA.
 AC Q9BHQ8;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
 GN AMA1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIGERIAN 050;
 RA Polley S.D., Conway D.J.;
 RT "Diversifying selection in domains of Plasmodium falciparum Apical
 Membrane Antigen 1 (AMA1).";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ408317; CAC34758.1; -.
 FT NON_TER 1
 FT NON_TER 437 437
 SQ SEQUENCE 437 AA; 50144 MW; 991DB6652BD8134F CRC64;

Query Match 95.0%; Score 95; DB 5; Length 437;
 Best Local Similarity 94.4%; Pred. No. 6.3e-08;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DGNCEIDPHVNEFSAIDL 18
 |||
 Db 171 DGNCEIDPHVNEFPAIDL 188

Search completed: January 29, 2002, 11:12:13
 Job time: 769 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:21:48 ; Search time 310.82 Seconds
(without alignments)
4.290 Million cell updates/sec

Title: US-09-763-397A-17

Perfect score: 102

Sequence: 1 GNAEKYDKMDEPQHGYKS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq.1101.*
1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	100.0	18	21	Plasmodium falcipa
2	102	100.0	350	21	Recombinant vaccin
3	102	100.0	622	10	Rhoptry membrane a
4	93	91.2	622	13	Plasmodium falcipa
5	93	91.2	622	16	Plasmodium falcipa
6	59	57.8	489	12	Plasmodium vivax 6
7	56	54.9	563	12	Plasmodium knowles
8	45	44.1	16	21	T cell antigen rec
9	43	42.2	180	21	Novel human G-prot
10	43	42.2	196	21	Novel human G-prot
11	43	42.2	508	21	Plasmodium falcipa

12	42	41.2	134	22	AAG82140	S. epidermidis ope
13	42	41.2	592	20	AAW97359	A 6-Transmembrane
14	41	40.2	187	21	AAW25391	Pinus radiata cell
15	41	40.2	294	20	AAW35663	Chlamydia pneumonia
16	41	40.2	561	10	AAW91588	Rhoptry membrane a
17	40	39.2	17	21	AAW66905	T cell antigen rec
18	40	39.2	344	19	AAW69762	Acetobacter xylinu
19	40	39.2	369	15	AAW59094	Murine IL-2R gamma
20	40	39.2	697	18	AAW32324	Staphylococcus aur
21	39.5	38.7	308	22	AAW81742	S. epidermidis ope
22	39	38.2	38	22	AAW75314	Human colon cancer
23	39	38.2	115	21	AAW24716	Plant SDF encoded
24	39	38.2	115	21	AAW24773	Plant SDF encoded
25	39	38.2	115	21	AAW24959	Plant SDF encoded
26	39	38.2	116	21	AAW61716	Arabidopsis thalia
27	39	38.2	155	21	AAW24715	Plant SDF encoded
28	39	38.2	155	21	AAW24958	Plant SDF encoded
29	39	38.2	170	21	AAW24957	Plant SDF encoded
30	39	38.2	172	21	AAW24957	Plant SDF encoded
31	39	38.2	228	21	AAW54065	Enzyme EPS1 which
32	39	38.2	228	21	AAW43767	Amino acid sequenc
33	39	38.2	279	17	AAW02367	Proteinase K varia
34	39	38.2	279	17	AAW02359	Proteinase K varia
35	39	38.2	334	18	AAW14070	S. thermophilus exo
36	39	38.2	404	16	AAW78621	Chicken GalNAc-alp
37	39	38.2	437	20	AAW99655	Human adenosine nu
38	39	38.2	484	18	AAW22169	S. thermophilus exo
39	39	38.2	497	22	AAW79105	Corynebacterium gl
40	39	38.2	511	22	AAW90799	C glutamicum prote
41	39	38.2	584	16	AAW75387	Natural resistance
42	38.5	37.7	3472	21	AAW90913	Cenarchaeum symbio
43	38	37.3	30	21	AAW15472	TCR beta V-N-J reg
44	38	37.3	31	21	AAW15474	TCR beta V-N-J reg
45	38	37.3	155	20	AAW00065	Enterococcus faeca

ALIGNMENTS

RESULT 1
AAW70293
ID AAW70293 standard; peptide: 18 AA.
XX
AC AAW70293;
XX
DT 06-JUN-2000 (first entry)
XX
DE Plasmodium falciparum AMA-1 antigenic epitope, P601.
XX
KW Recombinant protein; CDC/NIIMALVAC-1; multivalent; malaria; vaccine;
KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
KW Circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
KW Pfg27; antiparasitic; prevention; anti-CDC/NIIMALVAC-1 antibody.
XX
OS Plasmodium falciparum.
XX
PN WO200011179-A1.
XX
PD 02-MAR-2000.
XX
PF 19-AUG-1999; 99WO-US18869.
XX
PR 21-AUG-1998; 98US-0097703.
XX
PA (NATM-) NAT INST IMMUNOLOGY.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Lal AA, Shi YP, Hasnain SE;
XX
DR WPI; 2000-237654/20.

XX Novel recombinant protein as vaccine for treating malarial infection
 PT comprises antigenic peptides obtained from different stages of
 PT Plasmodium falciparum life cycle
 XX
 XX Claim 2; Page 16; 52pp; English.
 XX
 CC The present sequence is the antigenic epitope P601, derived from
 CC apical membrane antigen-1 (AMA-1) of the asexual blood stage of
 CC Plasmodium falciparum. It is used in the construction of recombinant
 CC protein CDC/NIIMALVAC-1, which is a multivalent, multistage malarial
 CC vaccine. The recombinant protein comprises, melittin signal peptide,
 CC (His)6 tag, T-cell epitope from tetanus toxoid and 21 antigenic epitopes
 CC from circumsporozoite protein (CSP), sporozoite surface protein-2
 CC (SSP-2), liver stage antigen-1 (LSA-1), merozoite surface protein-1
 CC (MSP-1), MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding
 CC antigen-175 (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete
 CC specific antigen, Pf27. These epitopes were obtained at different stages
 CC of the life cycle of P. falciparum. CDC/NIIMALVAC-1 vaccine has
 CC antiparasitic activity and can be used for treatment and prevention of
 CC malarial infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for
 CC detecting P. falciparum in biological samples.
 XX
 SQ Sequence 18 AA;

Query Match 100.0%; Score 102; DB 21; Length 18;
 Best Local Similarity 100.0%; Pred. No. 2e-10;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDPEQHYGKS 18
 Db 1 gnaekydkmdpeqhygks 18

RESULT 2
 AAY70278
 ID AAY70278 standard; Protein; 350 AA.
 XX
 AC AAY70278;

DT 06-JUN-2000 (first entry)

DE Recombinant vaccine CDC/NIIMALVAC-1.

KW Recombinant protein; CDC/NIIMALVAC-1; multivalent; malaria; vaccine;
 KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
 KW Circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
 KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
 KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
 KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
 KW Pf27; antiparasitic; prevention; anti-CDC/NIIMALVAC-1 antibody;
 KW honey bee.

OS Chimeric - Apis sp.
 OS Chimeric - Clostridium tetani.
 OS Chimeric - Plasmodium falciparum.

XX Key Location/Qualifiers

FT Peptide 1..22
 FT /label= Melittin_signal_peptide
 FT /note= "Derived from Honey bee"
 FT 23..350

FT Protein
 FT /label= Mature_CDC/NIIMALVAC-1
 FT /note= "Recombinant multivalent malarial vaccine"

XX WO200011179-A1.

PN 02-MAR-2000.

XX 19-AUG-1999; 99WO-0518869.

XX 21-AUG-1998; 98US-0097703.

XX
 PA (NAIM-) NAT INST IMMUNOLOGY.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Lal AA, Shi YP, Hasnain SE;
 XX
 DR WPI: 2000-237654/20.
 DR N-PSDB; AAZ51336.
 XX
 PT Novel recombinant protein as vaccine for treating malarial infection
 PT comprises antigenic peptides obtained from different stages of
 PT Plasmodium falciparum life cycle
 XX
 PS Claim 3; Page 43-44; 52pp; English.
 XX
 CC The present sequence is that of recombinant protein CDC/NIIMALVAC-1,
 CC which is a multivalent, multistage malarial vaccine. The recombinant
 CC protein comprises, melittin signal peptide, (His)6 tag, T-cell epitope
 CC from tetanus toxoid and 21 antigenic epitopes from circumsporozoite
 CC protein (CSP), sporozoite surface protein-2 (SSP-2), liver stage
 CC antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1), MSP-2, apical
 CC membrane antigen-1 (AMA-1), erythrocyte binding antigen-175 (EBA-175),
 CC rhoptry associated protein-1 (RAP-1) and gamete specific antigen, Pf27.
 CC These epitopes were obtained at different stages of the life cycle of
 CC Plasmodium falciparum. CDC/NIIMALVAC-1 vaccine has antiparasitic
 CC activity and can be used for treatment and prevention of malarial
 CC infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for detecting
 CC P. falciparum in biological samples.
 XX
 SQ Sequence 350 AA;

Query Match 100.0%; Score 102; DB 21; Length 350;
 Best Local Similarity 100.0%; Pred. No. 5.3e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDPEQHYGKS 18
 Db 158 gnaekydkmdpeqhygks 175

RESULT 3

AAP91632
 ID AAP91632 standard; protein; 622 AA.

XX AAP91632;

XX 25-JAN-1990 (first entry)

DT Rhoptry membrane antigen-1.

XX Malaria; rhoptry membrane antigen-1; antibodies.

XX Plasmodium falciparum D10.

XX WO8907645-A.

XX 24-AUG-1989.

XX 10-FEB-1989; 89WO-AU00056.

XX 12-FEB-1988; 88AU-0006743.

XX (SARA) SARAMANE PTY LTD.

XX Peterson MG, Crewther PE, Smythe JA, Marshall VM, Silva A;

XX WPI: 1989-263714/36.

XX N-PSDB; AAN90703.

XX Rhoptry membrane antigen of Plasmodium falciparum
 PT - used for producing antibodies and in immunisation,
 PT diagnostic and treatment methods for malaria.

XX PS Claim 1: Fig 3: 46pp; English.

XX CC RMA-1 can generate an immune response to malaria, and antibodies which
 CC can inhibit growth of the parasite. RMA-1 initially has mol. wt. 80 kD.
 XX SQ Sequence 622 AA;

Query Match 100.0%; Score 102; DB 10; Length 622;
 Best Local Similarity 100.0%; Pred. No. 9,9e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GNAEKYDKMDEPQHYGKS 18
 |||||
 Db 571 gnaekydkmdepqdygks 588

RESULT 4
 AAR27532
 ID AAR27532 standard; Protein; 622 AA.

XX AC AAR27532;
 XX DT 08-MAR-1993 (first entry)
 XX DE Plasmodium falciparum AMA1 antigen.

XX KW Recombinant poxvirus; antimalarial vaccine; malaria; immunise;
 immunogen.

XX OS Plasmodium falciparum.

XX PN WO9216616-A.

XX PD 01-OCT-1992.

XX PF 19-MAR-1992; 92WO-US02207.

XX PR 20-MAR-1991; 91US-0672183.

XX PR 18-MAR-1992; 92US-0852305.

XX PA (VIRO-) VIROGENETICS CORP.

XX PI De Taisne C, Paoletti E, Tine JA;

XX DR WPI: 1992-349203/42.

XX DR N-PSDB; AAQ29189.

XX PT Recombinant poxvirus - contg. Plasmodium DNA, useful as
 antimalarial vaccine

XX PS Example 4; Fig 5: 74pp; English.

XX CC This sequence is the Plasmodium falciparum AMA1 antigen.
 CC cDNA encoding it was cloned into vaccinia donor plasmids
 CC before being inserted into the vaccinia virus to be used in a
 CC vaccine to stimulate an antimalarial immunological response, or for
 CC in vitro prodn. of gene prods. for use as immunogens. As plasmodium
 CC genes are conserved among P. falciparum strains, they are widely
 CC effective in a vaccine.

XX SQ Sequence 622 AA;

Query Match 91.2%; Score 93; DB 13; Length 622;
 Best Local Similarity 94.4%; Pred. No. 3.1e-07;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPQHYGKS 18

Db 571 gnaekydkmdepqdygks 588

RESULT 5

AAR68840

ID AAR68840 standard; Protein; 622 AA.

XX AC AAR68840;

XX DT 24-AUG-1995 (first entry)

XX DE Plasmodium falciparum AMA-1 gene protein.

XX KW Plasmodium falciparum AMA-1 gene; recombinant poxvirus;
 KW multicomponent multistage malarial vaccines; immunogens;
 KW malaria diagnosis.

XX OS Plasmodium falciparum (3D7).

XX PN WO9428930-A.

XX PD 22-DEC-1994.

XX PF 10-JUN-1994; 94WO-US06652.

XX PR 11-JUN-1993; 93US-0075783.

XX PR 09-JUN-1994; 94US-0257073.

XX PA (VIRO-) VIROGENETICS CORP.

XX PI De Taisne C, Paoletti E, Tine JA;

XX DR WPI: 1995-036113/05.

XX DR N-PSDB; AAQ80910.

XX PT Recombinant poxvirus contg. Plasmodium DNA in non-essential
 PT region - useful in vaccines against malaria and for prodn. of

XX PT Plasmodium immunogens

XX PS Claim 3; Fig 5; 183pp; English.

XX CC AAQ80910 encodes AAR68840 the P. falciparum AMA-1 gene product. New
 CC recombinant poxviruses containing either the SERA, ABRA, Pfhap70,
 CC AMA-1, Pfs25, Pfs16, CSP, PfESP2, LSA-1, LSA-1 repeats, MSA-1,
 CC MSA-1 (N-terminal p83 or C-terminal gp42) genes, or a combination
 CC of these in non-essential regions of their genomes are claimed.
 CC These poxviruses (pref. with a virulence reducing genomic
 CC deletion or disruption) can be used as vaccines against malaria
 CC and for the prodn. of Plasmodium immunogens. These viruses
 CC provide multicomponent, multistage vaccines due to their expression
 CC of sporozoite, liver stage, blood stage and sexual stage proteins.

XX SQ Sequence 622 AA;

Query Match 91.2%; Score 93; DB 16; Length 622;
 Best Local Similarity 94.4%; Pred. No. 3.1e-07;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPQHYGKS 18

Db 571 gnaekydkmdepqdygks 588

RESULT 6

AAR10935

ID AAR10935 standard; Protein; 489 AA.

XX AC AAR10935;

XX DT 15-APR-1991 (first entry)

XX DE Plasmodium vivax 66kD merozoite antigen protein partial sequence.

XX KW Plasmodium merozoite antigens; malaria vaccine.

SQ sequence 16 AA;

Query Match 44.1%; Score 45; DB 21; Length 16;
Best Local Similarity 46.2%; Pred. No. 0.58;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 AEYDKMDEPQHY 15
| | | : : : | | |
Db 4 assydslnqpqh 16

RESULT 9

AA93140
ID AAY93140 standard; Protein; 180 AA.

XX AAY93140;

DT 06-DEC-2000 (first entry)

XX Novel human G-protein coupled receptor #1.

DE G-protein coupled receptor; human; bovine; nervous system disorder;

KW rat; mouse; somatostatin excretion.

XX Homo sapiens.

OS WO200029441-A1.

PN 25-MAY-2000.

XX 11-NOV-1999; 99WO-JP06283.

PR 13-NOV-1998; 98JP-0323759.

PR 08-MAR-1999; 99JP-0060030.

PR 14-APR-1999; 99JP-0106812.

PR 14-JUN-1999; 99JP-0166672.

PR 04-AUG-1999; 99JP-0221640.

PR 14-SEP-1999; 99JP-0259818.

XX (TAKE) TAKEDA CHEM IND LTD.

PA Watanabe T, Kikuchi K, Terao Y, Shintani Y, Hinuma S, Fukusumi S;

XX Fujii R, Hosoya M, Kitada C;

PI WPI: 2000-387747/33.

DR N-PSDB; AAA70500.

XX G protein coupled receptor protein and antibodies to it for treatment

PT and diagnosis of nerve diseases

PS Claim 1; Fig 1; 184pp; Japanese.

XX The invention relates to the isolation of novel G-protein coupled

CC receptor (GPCR) genes and their encoded proteins. This sequence

CC represents the protein sequence of a human GPCR. The DNAs and proteins

CC are used for the treatment, prevention and diagnosis of disorders of

CC the nervous system. The proteins and its fragments are also promoters

CC of somatostatin excretion.

XX Sequence 180 AA;

SQ

Query Match 42.2%; Score 43; DB 21; Length 180;
Best Local Similarity 57.1%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 EYDKMDEPQHYGK 17
| | | | | : : : | | |
Db 39 enydkyseprgypk 52

RESULT 10

AA93141

ID AAY93141 standard; Protein; 196 AA.

XX AAY93141;

DT 06-DEC-2000 (first entry)

XX Novel human G-protein coupled receptor #2.

DE G-protein coupled receptor; human; bovine; nervous system disorder;

KW rat; mouse; somatostatin excretion.

XX Homo sapiens.

PN WO200029441-A1.

XX 25-MAY-2000.

DT 11-NOV-1999; 99WO-JP06283.

PR 13-NOV-1998; 98JP-0323759.

PR 08-MAR-1999; 99JP-0060030.

PR 14-APR-1999; 99JP-0106812.

PR 14-JUN-1999; 99JP-0166672.

PR 04-AUG-1999; 99JP-0221640.

PR 14-SEP-1999; 99JP-0259818.

XX (TAKE) TAKEDA CHEM IND LTD.

PA Watanabe T, Kikuchi K, Terao Y, Shintani Y, Hinuma S, Fukusumi S;

XX Fujii R, Hosoya M, Kitada C;

PI WPI: 2000-387747/33.

DR N-PSDB; AAA70501.

XX G protein coupled receptor protein and antibodies to it for treatment

PT and diagnosis of nerve diseases

PS Claim 2; Fig 3; 184pp; Japanese.

XX The invention relates to the isolation of novel G-protein coupled

CC receptor (GPCR) genes and their encoded proteins. This sequence

CC represents the protein sequence of a human GPCR. The DNAs and proteins

CC are used for the treatment, prevention and diagnosis of disorders of

CC the nervous system. The proteins and its fragments are also promoters

CC of somatostatin excretion.

XX Sequence 196 AA;

SQ

Query Match 42.2%; Score 43; DB 21; Length 196;
Best Local Similarity 57.1%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 EYDKMDEPQHYGK 17
| | | | | : : : | | |
Db 39 enydkyseprgypk 52

RESULT 11

AA93141
ID AAB18187 standard; Protein; 508 AA.

XX AAB18187;

DT 07-NOV-2000 (first entry)

XX Plasmodium falciparum chromosome 2 related protein SEQ ID NO:44.

DE Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;

KW antimalarial; malaria; protozoicide; infection; insecticide.

XX Plasmodium falciparum.

OS


```

XX OS Homo sapiens.
XX PN EP897979-A2.
XX XX
XX PD 24-FEB-1999.
XX PF 06-AUG-1998; 98EP-0306292.
XX PR 16-DEC-1997; 97US-0991813.
XX PR 12-AUG-1997; 97US-0055375.
XX PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX PI Albone EF, Kikly KK;
XX DR WPI; 1999-134642/12.
XX DR N-PSDB; AAX15876.
XX PT New 6-Transmembrane protein (SDR2) polypeptide and polynucleotide -
XX PT useful as diagnostic reagents and for prevention and treatment of
XX PT cancer, AIDS and microbial infections
XX PS Claim 2; Page 6; 19pp; English.
XX CC The present sequence represents a 6-Transmembrane protein polypeptide
XX CC designated SDR2. SDR2 polypeptides and polynucleotides are useful for
XX CC diagnosing susceptibility to diseases by detecting mutations in the SDR2
XX CC gene, and can diagnose diseases associated with SDR2 protein imbalance by
XX CC determining SDR2 polypeptide expression levels. SDR2 polypeptides can be
XX CC used to screen for agonists and antagonists and to produce antibodies.
XX CC Diseases diagnosed, prevented and treated include: cancer, inflammation,
XX CC autoimmunity, allergy, asthma, rheumatoid arthritis, CNS inflammation,
XX CC cerebellar degeneration, Alzheimer's disease, Parkinson's disease,
XX CC multiple sclerosis, amyotrophic lateral sclerosis, head injury damage,
XX CC and other neurological abnormalities, septic shock, stroke,
XX CC osteoporosis, osteoarthritis, ischaemia reperfusion injury,
XX CC cardiovascular disease, kidney disease, liver disease, ischaemic injury,
XX CC myocardial infarction, hypotension, hypertension, AIDS, myelodysplastic
XX CC syndromes and other hematological abnormalities, aplastic anaemia, male
XX CC pattern baldness, and bacterial, fungal, protozoan and viral infections.
XX SQ Sequence 592 AA;

Query Match 41.2%; Score 42; DB 20; Length 592;
Best Local Similarity 53.3%; Pred. No. 96;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Oy 4 EKYDKMDEPOHYGKS 18
||||| |::| |
Db 350 ekydvtdspkniggs 364

RESULT 14
AAB25391
ID AAB25391 standard; Protein; 187 AA.
XX AC AAB25391;
XX DT 27-NOV-2000 (first entry)
XX DE Pinus radiata cell signalling involved protein SEQ ID NO:710.
XX KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
XX KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
XX KW environmental change; development; cell proliferation; differentiation;
XX KW elongation; survival; disease resistance; nutrient metabolism.
XX OS Pinus radiata.
XX PN WO2000042171-A1.
XX XX

```

```

PD 20-JUL-2000.
XX 11-JAN-2000; 2000WO-US00724.
XX PR 12-JAN-1999; 99US-0228986.
XX PR 01-NOV-1999; 99US-0162866.
XX PA (GENE-) GENESIS RES & DEV CORP LTD.
XX PI Strabala TJ, Nieuwenhuizen NJ;
XX DR WPI; 2000-476052/41.
XX PT Isolated polynucleotide encoding a polypeptide involved in cell
XX PT signalling used for generating transgenic plants with modified responses
XX PT to external signals -
XX PS Claim 3; Page 326-327; 527pp; English.
XX CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
XX CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
XX CC pine (Pinus radiata also known as Monterey pine). The protein sequences
XX CC are involved in cell signalling. The polynucleotide and protein
XX CC sequences can be used to modify the response of plant cells to external
XX CC signals e.g. environmental changes or pathogens during the growth and
XX CC development of a plant. They can be used to modify cell proliferation,
XX CC differentiation, elongation and survival, resistance to disease and
XX CC nutrient metabolism. Examples of modifications which can be produced are
XX CC altered fruit ripening and senescence of leaves and flowers e.g. to
XX CC delay senescence and prolong the life of cut flowers or enhance
XX CC senescence of reproductive organs to engineer sterile plants. Other
XX CC modifications can be used to delay senescence in selected cell types or
XX CC organs providing fruit and vegetables which have a longer shelf life
XX CC between harvest and consumption, or to decrease branching frequency in
XX CC forest tree species giving long stretches of valuable knot-free clear
XX CC wood which can be used in solid timber furniture and veneers.
XX SQ Sequence 187 AA;

Query Match 40.2%; Score 41; DB 21; Length 187;
Best Local Similarity 46.2%; Pred. No. 40;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Oy 5 KYDKMDEPOHYGK 17
||||| |::| |
Db 124 kydqiadpbfhgk 136

RESULT 15
AAY35663
ID AAY35663 standard; Protein; 294 AA.
XX AC AAY35663;
XX DT 13-SEP-1999 (first entry)
XX DE Chlamydia pneumoniae transmembrane protein sequence.
XX KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
XX KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
XX KW vaccine; neutralising epitope.
XX OS Chlamydia pneumoniae.
XX PN WO9927105-A2.
XX PD 03-JUN-1999.
XX PF 20-NOV-1998; 98WO-IB01890.
XX PR 04-NOV-1998; 98US-0107078.
XX PR 21-NOV-1997; 97FR-0014673.

```

```
XX (GEST.) GENSET.
XX PA
XX PI
XX Griffais R;
XX WPI; 1999-357842/30.
XX Genome sequence of Chlamydia pneumoniae
XX PT
XX PS Page 1382; Disclosure; 1912pp; English.
XX
XX AAY34584-Y35879 represent the proteins encoded by all the open reading
XX frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.
XX C. pneumoniae causes respiratory disease such as pneumonia and
XX bronchitis and is thought to be a contributing factor in heart
XX disease, sarcoidosis, sinusitis, purulent otitis media, erythema
XX nodosum or pharyngitis. The polypeptides encoded by the open reading
XX frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
XX immunogenic compositions as vaccines. Vectors containing C. pneumoniae
XX nucleotide sequences can also be used as immunogenic compositions,
XX especially where the vector directs the expression of a neutralising
XX epitope of C. pneumoniae.
XX SQ Sequence 294 AA;

Query Match 40.2%; Score 41; DB 20; Length 294;
Best Local Similarity 53.8%; Pred. No. 65;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 ERYDKMDEPOHYG 16
Db :||| : |||
91 kkydaiaapfhyg 103

Search completed: January 29, 2002, 10:21:49
Job time: 425 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:24:08 ; Search time 133.18 seconds
(without alignments)
3.041 Million cell updates/sec

Title: us-09-763-397a-17

Perfect score: 102

Sequence: 1 GNAEKYDKMDEPOHYGKS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	91.2	628	1	US-08-257-073-9
2	42	41.2	592	3	US-08-991-813-2
3	40	39.2	369	2	US-08-424-224-2
4	40	39.2	369	5	PCT-US94-02891-69
5	39	38.2	90	1	US-08-341-219-19
6	39	38.2	90	4	US-08-912-314A-19
7	39	38.2	404	2	US-08-666-367B-7
8	39	38.2	404	4	US-09-143-438-7
9	39	38.2	484	1	US-08-597-236-2
10	39	38.2	484	1	US-08-746-682A-2
11	39	38.2	584	4	US-08-637-823B-28
12	38	37.3	355	2	US-08-440-845D-25
13	38	37.3	441	1	US-08-121-713D-56
14	38	37.3	441	1	US-08-835-268-56
15	38	37.3	441	2	US-09-060-692-56
16	38	37.3	441	3	US-08-833-391-56
17	38	37.3	441	5	PCT-US94-10151A-56
18	38	37.3	7257	3	US-09-335-409-5
19	37.5	36.8	65	1	US-08-358-160-96
20	37.5	36.8	475	4	US-08-274-121B-4
21	37.5	36.8	476	2	US-08-569-150A-3
22	37	36.3	339	1	US-08-431-080-18
23	37	36.3	339	2	US-08-938-534-18
24	37	36.3	544	2	US-08-467-822-33
25	37	36.3	544	4	US-08-432-697-33
26	37	36.3	544	4	US-08-466-248-33
27	37	36.3	2259	4	US-09-413-814-70

28 37 36.3 2439 3 US-09-335-409-7 Sequence 7, Appli
29 37 36.3 2516 3 US-08-374-077C-2 Sequence 2, Appli
30 37 36.3 2516 4 US-08-895-590-2 Sequence 2, Appli
31 36 35.3 18 2 US-08-649-991-41 Sequence 41, Appl
32 36 35.3 132 2 US-08-649-991-65 Sequence 65, Appl
33 36 35.3 361 3 US-09-032-372-12 Sequence 12, Appl
34 36 35.3 409 4 US-09-140-466-5 Sequence 5, Appli
35 36 35.3 412 1 US-08-313-288B-18 Sequence 18, Appli
36 36 35.3 558 4 US-08-836-567-6 Sequence 6, Appli
37 36 35.3 681 2 US-08-272-255-6 Sequence 6, Appli
38 36 35.3 681 2 US-08-272-255-9 Sequence 9, Appli
39 36 35.3 681 3 US-08-964-288-5 Sequence 5, Appli
40 36 35.3 681 5 PCT-US95-08565-6 Sequence 6, Appli
41 36 35.3 681 5 PCT-US95-08565-9 Sequence 8, Appli
42 36 35.3 767 4 US-08-836-567-8 Sequence 8, Appli
43 35.5 34.8 102 1 US-08-202-389-27 Sequence 27, Appl
44 35.5 34.8 531 5 PCT-US92-00282-6 Sequence 6, Appli
45 35.5 34.8 593 1 US-08-202-389-12 Sequence 12, Appli

ALIGNMENTS

RESULT 1
US-08-257-073-9
; Sequence 9, Application US/08257073
; Patent No. 5766597
; GENERAL INFORMATION:
; APPLICANT: Paolletti, Enzo
; APPLICANT: de Taisne, Charles
; APPLICANT: Tine, John A.
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue, 25th Floor
; CITY: New York
; STATE: New York
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,073
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/075,783
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/852,305
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,183
; FILING DATE: 20-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2570
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 628 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

```
; FRAGMENT TYPE: internal
US-08-257-073-9

Query Match          91.2%; Score 93; DB 1; Length 628;
Best Local Similarity 94.4%; Pred. No. 1.9e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPOHYGKS 18
Db 576 GNAEKYDKMDEPODYGKS 593

RESULT 2
US-08-991-813-2
; Sequence 2, Application US/08991813
; Patent No. 6090579
; GENERAL INFORMATION:
; APPLICANT: ALBONE, EARL
; APPLICANT: KIKLY, KRISTINE
; TITLE OF INVENTION: HUMAN SDR2 CDNA CLONE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,813
; FILING DATE: 16-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/055,375
; FILING DATE: 12-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70281
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 592 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-991-813-2

Query Match          41.2%; Score 42; DB 3; Length 592;
Best Local Similarity 53.3%; Pred. No. 39;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 EKVDKMDPEOYHGS 18
Db 350 EKVDVTDSPKNGGS 364

RESULT 3
US-08-424-224-2
; Sequence 2, Application US/08424224
; Patent No. 5912173
; GENERAL INFORMATION:
; APPLICANT: LEONARD, WARREN J.
; TITLE OF INVENTION: MURINE IL-2R CDNA AND
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVE.
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT # 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,224
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/121,435
; FILING DATE: 14-SEPT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4061051
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 369
; TYPE: AMINO ACID
; TOPOLOGY: UNKNOWN
; MOLECULE TYPE:
; DESCRIPTION: PROTEIN
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: MURINE
; INDIVIDUAL ISOLATE: IL-2R
US-08-424-224-2

Query Match          39.2%; Score 40; DB 2; Length 369;
Best Local Similarity 31.2%; Pred. No. 50;
Matches 5; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPOHYG 16
Db 233 GSSQOWSKWSQPVHWG 248

RESULT 4
PCI-US94-02891-69
; Sequence 69, Application PC/TUS9402891
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN
; APPLICANT: SERVICES
; APPLICANT: OFFICE OF TECHNOLOGY TRANSFER, NATIONAL
; APPLICANT: INSTITUTE OF HEALTH, BOX OTT, BETHESDA, MARYLAND 20892 USA
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT OF
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVE.
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT # 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/02891
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/031,143
; FILING DATE: 12-MAR-1993
; APPLICATION NUMBER: 08/121,435
; FILING DATE: 14-SEPT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4061
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; TELEX: 421792
; SEQUENCE CHARACTERISTICS:
; LENGTH: 369
; TYPE: AMINO ACID
; TOPOLOGY: UNKNOWN
; MOLECULE TYPE: PROTEIN
; DESCRIPTION: PROTEIN
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: MURINE
; INDIVIDUAL ISOLATE: IL-2R
; PCT-US94-02891-69

Query Match 39.2%; Score 40; DB 5; Length 369;
Best Local Similarity 31.2%; Pred. No. 50;
Matches 5; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPHYG 16
I: : : : :
D 233 GSSQWQSKSQPVHWG 248

RESULT 5
US-08-341-219-19
; Sequence 19, Application US/08341219
; Patent No. 5643877
; GENERAL INFORMATION:
; APPLICANT: Zohar, Y.
; APPLICANT: Rivier, J.
; APPLICANT: Powell, J.
; APPLICANT: Sherwood, N.
; APPLICANT: Gothliff, Y.
; TITLE OF INVENTION: Compounds and Methods For Controlling
; TITLE OF INVENTION: Reproduction in Fish
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: N.Y.
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/341,219
; FILING DATE: 05-DEC-1994

; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi,, Laura A.
; REGISTRATION NUMBER: 30742
; REFERENCE/DOCKET NUMBER: 8399-003-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 90 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-341-219-19
Query Match 38.2%; Score 39; DB 1; Length 90;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 EKYDKMDEPHY 15
I: : : : :
D 48 KEVDQMAEPQHF 59

RESULT 6
US-08-912-314A-19
; Sequence 19, Application US/08912314A
; Patent No. 6210927
; GENERAL INFORMATION:
; APPLICANT: Zohar, Y.
; APPLICANT: Rivier, J.
; APPLICANT: Powell, J.
; APPLICANT: Sherwood, N.
; APPLICANT: Gothliff, Y.
; TITLE OF INVENTION: Compounds and Methods For Controlling
; TITLE OF INVENTION: Reproduction in Fish
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: N.Y.
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,314A
; FILING DATE: 30-JUN-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/341,219
; FILING DATE: 05-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi,, Laura A.
; REGISTRATION NUMBER: 30742
; REFERENCE/DOCKET NUMBER: 8399-003-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 90 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant

;
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-912-314A-19

Query Match 38.2%; Score 39; DB 4; Length 90;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 4 EKYDKMDEPQHY 15
Db 48 KEVDQMAEPQHF 59

RESULT 7
US-08-666-367B-7
; Sequence 7, Application US/08666367B
; Patent No. 5854042
; GENERAL INFORMATION:
; APPLICANT: Shuichi TSUJI et al.
; TITLE OF INVENTION: NOVEL SUGAR-CHAIN SYNTHETASE AND PROCESS FOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/666,367B
; FILING DATE: August 19, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:

TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 404 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: G. gallus (chicken)
; US-08-666-367B-7

Query Match 38.2%; Score 39; DB 2; Length 404;
Best Local Similarity 66.7%; Pred. No. 80;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 EKYDKMDEPQHY 15
Db 281 EGLDKGDEPQKY 292

RESULT 8

US-09-143-438-7
; Sequence 7, Application US/09143438
; Patent No. 6218161
; GENERAL INFORMATION:
; APPLICANT: Shuichi TSUJI et al.
; TITLE OF INVENTION: NOVEL SUGAR-CHAIN SYNTHETASE AND PROCESS FOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., #800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/143,438
; FILING DATE: August 28, 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/666,367
; FILING DATE: August 19, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-721-8200
; TELEFAX: 202-721-8250
; TELEX:

INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 404 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: G. gallus (chicken)
; US-09-143-438-7

Query Match 38.2%; Score 39; DB 4; Length 404;
Best Local Similarity 66.7%; Pred. No. 80;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 EKYDKMDEPQHY 15
Db 281 EGLDKGDEPQKY 292

RESULT 9
US-08-597-236-2
; Sequence 2, Application US/08597236
; Patent No. 5733765
; GENERAL INFORMATION:
; APPLICANT: STINGELE, Francesca
; APPLICANT: MOLLET, Beat

; TITLE OF INVENTION: LACTIC BACTERIA PRODUCING
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: pennie & Edmonds
; STREET: 1155 Avenue of the Americans
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/597,236
FILING DATE:
CLASSIFICATION: 426
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95201669.9
FILING DATE: 20-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci A., Allan
REGISTRATION NUMBER: 30256
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 484 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-597-236-2

Query Match 38.2%; Score 39; DB 1; Length 484;
Best Local Similarity 43.8%; Pred. No. 97;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPQHYG 16
| | : | | | : | |
Db 280 GGADQYDKLTHAGIYG 295

RESULT 10
US-08-746-682A-2
Sequence 2, Application US/08746682A
Patent No. 5786184
GENERAL INFORMATION:
APPLICANT: STINGELE, Franscesca
APPLICANT: MOLLET, Beat
TITLE OF INVENTION: LACTIC BACTERIA PRODUCING
EXOPOLYSACCHARIDES
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americans
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/746,682A
FILING DATE: 14-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/597,236
FILING DATE: 20-JUN-1995
APPLICATION NUMBER: EP 95201669.9
FILING DATE: 20-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci A., Allan
REGISTRATION NUMBER: 30256
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 484 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-746-682A-2

Query Match 38.2%; Score 39; DB 1; Length 484;
Best Local Similarity 43.8%; Pred. No. 97;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPQHYG 16
| | : | | | : | |
Db 280 GGADQYDKLTHAGIYG 295

RESULT 11
US-08-637-823B-28
Sequence 28, Application US/08637823B
Patent No. 6184031
GENERAL INFORMATION:
APPLICANT: Gros, Philippe
APPLICANT: Skamene, Emil
TITLE OF INVENTION: DNA SEQUENCES THAT ENCODE A NATURAL
RESISTANCE TO INFECTION WITH INTRACELLULAR PARASITES
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: KLAUBER & JACKSON
STREET: 411 Hackensack Ave
CITY: Hackensack
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/637,823B
FILING DATE: 05/08/96
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson, David A.
REGISTRATION NUMBER: 26,742
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487 5800
TELEFAX: 201 343 1684
TELEX:
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 584 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORGANISM: Mus musculus
TISSUE TYPE: Blood
CELL TYPE: B-cell precursor
CELL LINE: 70/2
IMMEDIATE SOURCE:
CLONE: M9ramp-2
US-08-637-823B-28

Query Match 38.2%; Score 39; DB 4; Length 584;
Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPQ 13
| | | | | : | |
Db 1 GGAESYSKSTDPQ 13

RESULT 12
US-08-440-845D-25
Sequence 25, Application US/08440845D
Patent No. 5955329

GENERAL INFORMATION:
APPLICANT: Yuan, L.
APPLICANT: Kridl, J.
APPLICANT: Denesh, K.
APPLICANT: Knauf, V.
TITLE OF INVENTION: Engineering Plant Thioesterases For
ALTERED Substrate Specificity.
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.40 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,845D
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 113
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 residues
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
OTHER INFORMATION: Xaa = Lys or Pro
US-08-440-845D-25

Query Match 37.3% Score 38; DB 2; Length 355;
Best Local Similarity 42.9%; Pred. No. 1e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 4 EYDKMDEPQHYGK 17
: | : | : | : |
Db 229 KKISKLDPAQYSK 242

RESULT 13
US-08-121-713D-56
Sequence 56, Application US/08121713D
Patent No. 5639856
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Matthes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/121,713D
FILING DATE: 13-SEP-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 441 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-121-713D-56

Query Match 37.3% Score 38; DB 1; Length 441;
Best Local Similarity 42.1%; Pred. No. 1.3e+02;
Matches 8; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

Qy 1 GNABE-KYDKMDEPQHYGK 17
: | : | : | : | : | : |
Db 122 GNPCKWKIDGSDDPKHRGR 140

RESULT 14
US-08-835-268-56
Sequence 56, Application US/08835268
Patent No. 5807826
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Matthes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/835,268
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/121,713
FILING DATE: 13-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:

; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-835-268-56

Query Match 37.3%; Score 38; DB 1; Length 441;
Best Local Similarity 42.1%; Pred. No. 1.3e+02;
Matches 8; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

Qy 1 GNAE--KYDKWDEPOHYCK 17
||:|:|:|:|:
Db 122 GNPCKWKIDGSDDPKHRGR 140

RESULT 15

US-09-060-692-56
; Sequence 56, Application US/09060692
; Patent No. 5935865
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,692
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-060-692-56

Query Match 37.3%; Score 38; DB 2; Length 441;
Best Local Similarity 42.1%; Pred. No. 1.3e+02;
Matches 8; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

Qy 1 GNAE--KYDKWDEPOHYCK 17
||:|:|:|:|:
Db 122 GNPCKWKIDGSDDPKHRGR 140

Search completed: January 29, 2002, 10:24:09
Job time: 515 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:26:43 ; Search time 144.96 Seconds
(without alignments)
9.459 Million cell updates/sec

Title: US-09-763-397A-17

Perfect score: 102

Sequence: 1 GNAEKYDKMDEPQHYGKS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_68:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102	100.0	622	2 A32499	apical membrane an
2	99	97.1	622	2 C44986	apical membrane an
3	93	91.2	622	2 D4986	apical membrane an
4	93	91.2	622	2 B4986	apical membrane an
5	90	88.2	622	2 A44986	apical membrane an
6	56	54.9	563	2 A39238	66K merozoite surf
7	49	48.0	562	2 A44944	apical membrane an
8	44	43.1	371	2 T34410	hypothetical prote
9	43	42.2	508	2 E71620	hypothetical prote
10	42	41.2	188	2 B42983	orf 3'of dae - Lac
11	42	41.2	379	2 T40312	probable splicing
12	42	41.2	384	2 T29682	hypothetical prote
13	42	41.2	835	1 A49891	outer membrane ush
14	42	41.2	971	2 T24866	hypothetical prote
15	41	40.2	68	2 S60797	M protein precuro
16	41	40.2	91	1 C69973	ribonuclease inhib
17	41	40.2	266	2 A81531	methionine aminope
18	41	40.2	285	2 T41656	vacuolar ATP synth
19	41	40.2	291	2 F86616	methionine aminope
20	41	40.2	291	2 E72008	methionine aminope
21	41	40.2	294	2 D83136	hypothetical prote
22	41	40.2	384	1 S54484	probable fatty aci
23	41	40.2	502	2 JE0295	L1 protein - human
24	41	40.2	558	2 A44964	apical membrane an
25	41	40.2	712	2 B47021	pectic enzyme secr
26	41	40.2	813	2 D64527	hypothetical prote
27	41	40.2	1017	2 SG2435	probable glycine d
28	41	40.2	1110	2 T33877	hypothetical prote
29	41	40.2	1635	2 T32452	hypothetical prote

30 40.5 39.7 170 2 H70034
31 40.5 39.7 687 2 T27533
32 40.5 39.7 700 2 S09699
33 40.5 39.7 838 2 G81444
34 40 39.2 323 2 G81441
35 40 39.2 344 2 A43592
36 40 39.2 369 2 I49280
37 40 39.2 383 2 B82157
38 40 39.2 443 2 T17868
39 40 39.2 489 2 F72324
40 40 39.2 506 2 C64377
41 40 39.2 564 2 S09199
42 40 39.2 973 2 T40778
43 39 38.2 90 1 RHMSG
44 39 38.2 148 2 T19724
45 39 38.2 155 2 D86272

hypothetical prote
hypothetical prote
bib protein - frui
molybdopterin-cont
nifu protein homol
outer membrane pro
interleukin-2 rece
eha protein VC1798
hypothetical prote
sensor histidine k
hypothetical prote
cytochrome-c3 hydr
hypothetical 129.5
gonadoliberin prec
hypothetical prote
hypothetical prote

ALIGNMENTS

RESULT 1

A32499

apical membrane antigen 1 - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum.

C:Date: 12-Oct-1989 #sequence_revision 12-Oct-1989 #text_change 09-Jun-2000

C:Accession: A32499

R:Peterson, M.G.; Marshall, V.M.; Smythe, J.A.; Crewther, P.E.; Lew, A.; Silva, A.; A

Mol. Cell. Biol. 9, 3151-3154, 1989

A:Title: Integral membrane protein located in the apical complex of Plasmodium falcip

A:Reference number: A32499; MUID:89384584

A:Accession: A32499

A>Status: preliminary

A:Molecule type: DNA; mRNA

A:Residues: 1-622 <PEP>

A:Cross-references: GB:M27133; MID:g160072; PID:g160073

C:Keywords: membrane protein; surface antigen

Query Match 100.0%; Score 102; DB 2; Length 622;

Best Local Similarity 100.0%; Pred. No. 1.8e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GNAEKYDKMDEPQHYGKS 18

Db 571 GNAEKYDKMDEPQHYGKS 588

RESULT 2

C44986

apical membrane antigen 1 - malaria parasite (Plasmodium falciparum) (strain FCR 3)

C:Species: Plasmodium falciparum

C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000

C:Accession: C44986

R:Thomas, A.W.; Waters, A.P.; Carr, D.

Mol. Biochem. Parasitol. 42, 285-288, 1990

A:Title: Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate a

A:Reference number: A44986; MUID:91101665

A:Accession: C44986

A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: DNA

A:Residues: 1-622 <THO>

A:Cross-references: GB:M34554

C:Keywords: membrane protein; surface antigen

Query Match

Best Local Similarity 97.1%; Score 99; DB 2; Length 622;

Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GNAEKYDKMDEPQHYGKS 18

Db 571 GNAEKYDKMDEPQHYGKS 588

RESULT 3

D44986
apical membrane antigen 1 - malaria parasite (Plasmodium falciparum) (strain 768)
C:Species: Plasmodium falciparum
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000
C:Accession: D44986
R:Thomas, A.W.; Waters, A.P.; Carr, D.
Mol. Biochem. Parasitol. 42, 285-288, 1990
A:Title: Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate anti-
A:Accession: D44986
A:Reference number: A44986; MUID:91101665
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: DNA
A:Residues: 1-622 <THO>
A:Cross-references: GB:M34555
C:Keywords: membrane protein; surface antigen

Query Match 91.2%; Score 93; DB 2; Length 622;

Best Local Similarity 94.4%; Pred. No. 5e-07;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPQHYGKS 18

|||||

Db 571 GNAEKYDKMDEPQDYGKS 588

RESULT 4

B44986
apical membrane antigen 1 - malaria parasite (Plasmodium falciparum) (strain Thai Tn)
C:Species: Plasmodium falciparum
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000
C:Accession: B44986
R:Thomas, A.W.; Waters, A.P.; Carr, D.
Mol. Biochem. Parasitol. 42, 285-288, 1990
A:Title: Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate anti-
A:Reference number: A44986; MUID:91101665
A:Accession: B44986
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: DNA
A:Residues: 1-622 <THO>
A:Cross-references: GB:M34553
C:Keywords: membrane protein; surface antigen

Query Match 91.2%; Score 93; DB 2; Length 622;

Best Local Similarity 94.4%; Pred. No. 5e-07;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPQHYGKS 18

|||||

Db 571 GNAEKYDKMDEPQDYGKS 588

RESULT 5

A44986
apical membrane antigen 1 - malaria parasite (Plasmodium falciparum) (strain CAMP)
C:Species: Plasmodium falciparum
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000
C:Accession: A44986
R:Thomas, A.W.; Waters, A.P.; Carr, D.
Mol. Biochem. Parasitol. 42, 285-288, 1990
A:Title: Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate anti-
A:Reference number: A44986; MUID:91101665
A:Accession: A44986
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: DNA
A:Residues: 1-622 <THO>
A:Cross-references: GB:M34552
C:Keywords: membrane protein; surface antigen

Query Match 88.2%; Score 90; DB 2; Length 622;

Best Local Similarity 88.9%; Pred. No. 1.5e-06;

Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPQHYGKS 18

|||||

Db 571 GNAEKYDKMDEPQDYGKS 588

RESULT 6

A39238
66K merozoite surface antigen precursor - Plasmodium knowlesi
C:Species: Plasmodium knowlesi
C:Date: 13-Sep-1991 #sequence_revision 13-Sep-1991 #text_change 07-Feb-1997
C:Accession: A39238
R:Waters, A.P.; Thomas, A.W.; Deans, J.A.; Mitchell, G.H.; Hudson, D.E.; Miller, L.H.
J. Biol. Chem. 265, 17974-17979, 1990
A:Title: A merozoite receptor protein from Plasmodium knowlesi is highly conserved an
A:Reference number: A39238; MUID:91009268
A:Accession: A39238
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-563 <WAT>
A:Cross-references: GB:J05631
C:Keywords: surface antigen

Query Match 54.9%; Score 56; DB 2; Length 563;

Best Local Similarity 60.0%; Pred. No. 0.39;

Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 EKYDKMDEPQHYGKS 18

|||||

Db 514 DKYDKMDQAEAYGKT 528

RESULT 7

A44944
apical membrane antigen 1 precursor - Plasmodium fragile
C:Species: Plasmodium fragile
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-May-2000
C:Accession: A44944
R:Peterson, M.G.; Nguyen-Dinh, P.; Marshall, V.M.; Elliott, J.F.; Collins, W.E.; Ande
Mol. Biochem. Parasitol. 39, 279-284, 1990
A:Title: Apical membrane antigen of Plasmodium fragile.
A:Reference number: A44944; MUID:90205978
A:Accession: A44944
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-562 <PET>
A:Cross-references: GB:M29898

Query Match 48.0%; Score 49; DB 2; Length 562;

Best Local Similarity 57.1%; Pred. No. 5.1;

Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 EKYDKMDEPQHYGK 17

|||||

Db 514 DKYDKMEQADYGK 527

RESULT 8

T34410
hypothetical protein F07E5.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34410
R:Du, Z.; Goela, D.
submitted to the EMBL Data Library, December 1996
A:Description: The sequence of C. elegans cosmid F07E5.
A:Reference number: Z21520

A:Accession: T34410
A:Status: preliminary;
A:Molecule type: DNA
A:Residues: 1-371 <DUZ>
A:Cross-references: PIDN:AAB37904.1; G
A:Experimental source: strain Bristol N2; clone F07
C:Genetics:
A:Gene: CESP:F07E5.5
A:Map position: 2
A:Introns: 76/3; 134/3; 167/3; 227/2; 276/1; 293/1;

R:Schifferli, D.M.; Alrutz, M.A.
J. Bacteriol. 176, 1099-1110, 1994
A:Title: Permissive linker insertion sites in the outer membrane protein of 987P fimbriae
A:Reference number: A49891; MUID:94148769

A:Accession: A49891
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-835 <SCH>
A:Cross-references: GB:I22659; NID:9437334; PIDN:AAA21827.1; PID:9437336
C:Genetics:
A:Gene: fasD
C:Superfamily: outer membrane usher protein fimD
C:Keywords: membrane protein

Query Match 41.2%; Score 42; DB 1; Length 835;
Best Local Similarity 57.1%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 AEKYDKMDEPQHYG 16
| | | | | | | | | |
Db 425 ANKYDKIHSGQSYG 438

RESULT 14

T24866
hypothetical protein T12D8.9 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T24866; T25274
R:McMurray, A.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19944
A:Accession: T24866
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-971 <WIL>
A:Cross-references: EMBL:Z81120; PIDN:CAB03350.1; GSPDB:GN00021; CESP:T12D8.9
A:Experimental source: clone T12D8
R:Gardner, A.
submitted to the EMBL Data Library, December 1996
A:Reference number: Z20008
A:Accession: T25274
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-971 <WI2>
A:Cross-references: EMBL:Z83241; PIDN:CAB05819.1; GSPDB:GN00021; CESP:T12D8.9
A:Experimental source: clone T25C8
C:Genetics:
A:Gene: CESP:T12D8.9
A:Map position: 3
A:Introns: 148/2; 929/1

Query Match 41.2%; Score 42; DB 2; Length 971;
Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPQ 13
| | | | | | | | | |
Db 402 GSASKYDKVKTQ 414

RESULT 15

S60797
M protein precursor - *Streptococcus pyogenes* (serotype M17) (fragment)
C:Species: *Streptococcus pyogenes*
A:Variety: serotype M17
C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 26-Aug-1999
C:Accession: S60797
R:Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.
Mol. Microbiol. 14, 619-631, 1994
A:Title: Non-congruent relationships between variation in emm gene sequences and the pop

A:Reference number: S60784; MUID:95198537
A:Accession: S60797
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-68 <WHA>
A:Cross-references: EMBL:U11932; NID:g533541; PIDN:AAA99548.1; PID:gl235802
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Superfamily: M5 protein

Query Match 40.2%; Score 41; DB 2; Length 68;
Best Local Similarity 63.6%; Pred. No. 10;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 EKYDKMDEPQH 14
| | | | | | | | | |
Db 44 EKYDKIDITENH 54

Search completed: January 29, 2002, 10:26:44
Job time: 655 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:13:45 ; Search time 80.65 Seconds
(without alignments)
8.183 Million cell updates/sec

Title: US-09-763-397A-17

Perfect score: 102
Sequence: 1 GNAEKYDKMDEPOHYGKS 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	ID	Description
1	102	100.0	622	1 AMAL_PLAFF	P22621 plasmodium
2	99	97.1	622	1 AMAL_PLAFG	P50490 plasmodium
3	93	91.2	622	1 AMAL_PLAFB	P50492 plasmodium
4	93	91.2	622	1 AMAL_PLAFH	P50491 plasmodium
5	90	88.2	622	1 AMAL_PLAFC	P50489 plasmodium
6	56	54.9	563	1 PK66_PLAKU	P21303 plasmodium
7	49	48.0	562	1 AMAL_PLAFR	P22622 plasmodium
8	42	41.2	405	1 DITB_LACCA	P35855 lactobacill
9	42	41.2	835	1 FASD_ECOLI	P46000 escherichia
10	41	40.2	285	1 VATD_SCHPO	O59823 schizosacch
11	41	40.2	291	1 AMPM_CHLPN	O92690 chlamydia p
12	41	40.2	384	1 SC57_YEAST	O03529 saccharomyc
13	41	40.2	558	1 AMAL_PLACH	P16445 plasmodium
14	41	40.2	712	1 GSPD_ERWRCH	P31700 erwinia chr
15	41	40.2	1017	1 GSPD_SCHPO	O09785 schizosacch
16	40.5	39.7	700	1 BIB_DROME	P23645 drosophila
17	40	39.2	92	1 GONL_CAVPO	O54713 cavia porce
18	40	39.2	344	1 TMPL_TREPH	P29721 treponema p
19	40	39.2	369	1 CYRG_MOUSE	P34902 mus musculu
20	40	39.2	506	1 Y619_METJA	O58036 methanococc
21	40	39.2	563	1 PHNL_DESFR	P18188 desulfovibr
22	40	39.2	1105	1 DPOD_ORYSA	O91re6 oryza sativ
23	39	38.2	90	1 GONL_MOUSE	P13562 mus musculu
24	39	38.2	92	1 GONL_TUPGB	O95335 tupatia glis
25	39	38.2	251	1 TRPC_HALVO	P18304 halobacteri
26	39	38.2	312	1 RPOA_LACLA	O9cdy3 lactococcus
27	39	38.2	404	1 CAG5_CHICK	O92184 gallus gall
28	39	38.2	1023	1 TSQC_PSEAM	P55019 pseudopleur
29	38.5	37.7	191	1 ALLA_SCHPO	O13843 schizosacch
30	38.5	37.7	515	1 YH80_YEAST	O03264 saccharomyc
31	38	37.3	211	1 YG3K_YEAST	P48235 saccharomyc
32	38	37.3	373	1 CYRG_CANFA	P40321 canis famil
33	38	37.3	403	1 VA39_VACCV	P21062 vaccinia vi

RESULT 1

ID	AMAL_PLAFF	STANDARD;	PRT;	622 AA.
AC	P22621;			
DT	01-AUG-1991 (Rel. 19, Created)			
DT	01-AUG-1991 (Rel. 19, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).			
GN	AMA-1 OR PF83.			
OS	Plasmodium falciparum (isolate FC37 / Papua New Guinea).			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
OX	NCBI_TaxID=5837;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89384584; PubMed=2701947;			
RA	Peterson M.G., Marshall V.M., Smythe J.A., Crewther P.E., Lew A.,			
RA	Silva A., Anders R.F., Kemp D.J.			
RT	"Integral membrane protein located in the apical complex of			
RT	Plasmodium falciparum."			
RL	Mol. Cell. Biol. 9:3151-3154(1989).			
CC	-!- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.			
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	-!- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO			
CC	PK66 FROM P.KNOWLESI.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M27133; AAA29475.1; -			
DR	EMBL; M27957; AAA29476.1; -			
DR	EMBL; A08267; CAA00764.1; -			
DR	PIR; A32499; A32499.			
DR	InterPro; IPR003298; Apmem_Ag1.			
DR	Pfam; PF02430; AMA-1; 1.			
KW	Malaria; Signal; Transmembrane; Antigen; Glycoprotein.			
FT	SIGNAL 1 24 POTENTIAL.			
FT	CHAIN 1 25 622			
FT	DOMAIN 25 546			
FT	TRANSMEM 547 567			
FT	DOMAIN 568 622			
FT	DOMAIN 162 162			
FT	CARBOHYD 286 286			
FT	CARBOHYD 371 371			
FT	CARBOHYD 421 421			
FT	CARBOHYD 422 422			
FT	CARBOHYD 499 499			
FT	CONFLICT 308 308			
FT	CONFLICT 332 332			
FT	CONFLICT 407 407			
FT	CONFLICT 439 439			
FT	CONFLICT 496 496			

Q57706 methanococc
P21764 vaccinia vi
Q01532 saccharomyc
Q59714 pseudomonas
P94132 acinetobact
P27902 streptomyce
P31780 aeromonas h
P45778 aeromonas s
P41428 autographa
P42282 drosophila
Q09750 schizosacch
P53886 saccharomyc

ALIGNMENTS

```
FT CONFLICT 503 N -> R (IN REF. 1: AAA29476).
SQ SEQUENCE 622 AA; 72009 MW; 7D41335E249FA18F CRC64;

Query Match 100.0%; Score 102; DB 1; Length 622;
Best Local Similarity 100.0%; Pred. No. 7e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPQHYGKS 18
Db 571 GNAEKYDKMDEPQHYGKS 588
|||||

RESULT 2
AMAL_PLAF8
ID AMAL_PLAF8 STANDARD; PRT; 622 AA.
AC P50492;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
GN AMA-1 OR PF83.
OS Plasmodium falciparum (isolate FCR-3 / Gambia).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5838;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91101665; PubMed-2270110;
RA Thomas A.W., Waters A.P., Carr D.;
RT "Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate antigen of plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 42:285-287(1990).
CC -1- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO PK66 FROM P.KNOWLES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
DR EMBL; M58546; AAA29719.1; -
DR InterPro; IPR003298; Apmem_Ag1.
DR Pfam; PF02430; AMA-1; 1.
KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
FT CHAIN 1 24 POTENTIAL.
FT DOMAIN 25 546 APICAL MEMBRANE ANTIGEN 1.
FT TRANSMEM 547 567 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 422 422 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 622 AA; 71967 MW; 089336BE0464695C CRC64;
DR EMBL; M58546; AAA29719.1; -
DR InterPro; IPR003298; Apmem_Ag1.
DR Pfam; PF02430; AMA-1; 1.
KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
FT CHAIN 1 24 POTENTIAL.
FT DOMAIN 25 546 APICAL MEMBRANE ANTIGEN 1.
FT TRANSMEM 547 567 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 422 422 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 622 AA; 71967 MW; 089336BE0464695C CRC64;

Query Match 97.1%; Score 99; DB 1; Length 622;
Best Local Similarity 94.4%; Pred. No. 2.1e-08;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPQHYGKS 18
Db 571 GNAEKYDKMDEPQHYGKS 588
|||||

RESULT 3
AMAL_PLAF8
ID AMAL_PLAF8 STANDARD; PRT; 622 AA.
AC P50492;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
GN AMA-1 OR PF83.
OS Plasmodium falciparum (isolate FCR-3 / Gambia).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5838;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91101665; PubMed-2270110;
RA Thomas A.W., Waters A.P., Carr D.;
RT "Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate antigen of plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 42:285-287(1990).
CC -1- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO PK66 FROM P.KNOWLES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
DR EMBL; M58546; AAA29719.1; -
DR InterPro; IPR003298; Apmem_Ag1.
DR Pfam; PF02430; AMA-1; 1.
KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
FT CHAIN 1 24 POTENTIAL.
FT DOMAIN 25 546 APICAL MEMBRANE ANTIGEN 1.
FT TRANSMEM 547 567 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 422 422 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 622 AA; 71967 MW; 089336BE0464695C CRC64;

Query Match 91.2%; Score 93; DB 1; Length 622;
Best Local Similarity 94.4%; Pred. No. 2e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPQHYGKS 18
Db 571 GNAEKYDKMDEPQHYGKS 588
|||||

RESULT 4
AMAL_PLAF8
ID AMAL_PLAF8 STANDARD; PRT; 622 AA.
AC P50492;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
GN AMA-1 OR PF83.
OS Plasmodium falciparum (isolate thtn / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=70151;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91101665; PubMed-2270110;
```

RA Thomas A.W., Waters A.P., Carr D.;
RT "Analysis of variation in PF83, an erythrocytic merozoite vaccine
RL candidate antigen of Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 42:285-287(1990).
CC -!- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO
CC PK66 FROM P.KNOWLES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M58547; AAA29720.1; -;
DR InterPro; IPR003298; Apmem_Ag1.
DR Pfam; PF02430; AMA-1; 1.
KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 622 APICAL MEMBRANE ANTIGEN 1.
FT DOMAIN 25 546 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 547 567 POTENTIAL.
FT DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 422 422 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 622 AA; 71989 MW; 1FDFA53593C94CC5 CRC64;

Query Match 91.2%; Score 93; DB 1; Length 622;
Best Local Similarity 94.4%; Pred. No. 2e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GNAEKYDKMDEPQHYGKS 18
Db 571 GNAEKYDKMDEPQDYGKS 588
|||||
RESULT 5
ANAL_PLAFC
ID AMAL_PLAFC STANDARD; PRT; 622 AA.
AC P50489;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
GN AMA-1 OR PF83.
OS Plasmodium falciparum (isolate Camp / Malaysia).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5835;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91101665; PubMed=2270110;
RA Thomas A.W., Waters A.P., Carr D.;
RT "Analysis of variation in PF83, an erythrocytic merozoite vaccine
RL candidate antigen of Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 42:285-287(1990).
CC -!- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO
CC PK66 FROM P.KNOWLES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M58545; AAA29718.1; -;
DR InterPro; IPR003298; Apmem_Ag1.
DR Pfam; PF02430; AMA-1; 1.
KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 622 APICAL MEMBRANE ANTIGEN 1.
FT DOMAIN 25 546 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 547 567 POTENTIAL.
FT DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 422 422 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 622 AA; 71943 MW; 26CE8CF76D07C637 CRC64;

Query Match 88.2%; Score 90; DB 1; Length 622;
Best Local Similarity 88.9%; Pred. No. 6.1e-07;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GNAEKYDKMDEPQHYGKS 18
Db 571 GNAEKYDKMDEPQDYGKS 588
|||||
RESULT 6
PK66_PLAKU
ID PK66_PLAKU STANDARD; PRT; 563 AA.
AC P21303;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE MEROZOITE RECEPTOR PK66 PRECURSOR (66 KDA PROTECTIVE MINOR SURFACE
DE ANTIGEN).
GN PK66.
OS Plasmodium knowlesi (strain nuri).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5852;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LINE W1;
RX MEDLINE=91009268; PubMed=2211675;
RA Waters A.P., Thomas A.W., Deans J.A., Mitchell G.H., Hudson D.E.,
RA Miller L.H., McCutchan T.F., Cohen S.;
RT "A merozoite receptor protein from Plasmodium knowlesi is highly
RT conserved and distributed throughout Plasmodium.";
RL J. Biol. Chem. 265:17974-17979(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NURI;
RX MEDLINE=911817058; PubMed=2011149;
RA Waters A.P., Thomas A.W., Mitchell G.H., McCutchan T.F.;
RT "Intra-generic conservation and limited inter-strain variation in a
RT protective minor surface antigen of Plasmodium knowlesi merozoites.";
RL Mol. Biochem. Parasitol. 44:141-144(1991).
CC -!- FUNCTION: MEROZOITE RECEPTOR PK66 IS A SURFACE ANTIGEN INVOLVED
CC IN PARASITE INVASION OF ERYTHROCYTES.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CONCENTRATED AT THE
CC APICAL END PRIOR TO RUPTURE, FOLLOWING WHICH IT CAN DISTRIBUTE
CC ITSELF ENTIRELY ACROSS THE SURFACE OF THE FREE MEROZOITE. DURING
CC INVASION PK66 IS EXCLUDED FROM THE ERYTHROCYTE AT, AND BEHIND, THE
CC INVASION INTERFACE.
CC -!- MISCELLANEOUS: PK66 EXPRESSED IN MATURE SCHIZONTS IS RAPIDLY
CC PROCESSED AS THE SCHIZONT RUPTURES, YIELDING A 42/44 KDA DOUBLET
CC ASSOCIATED WITH THE SURFACE MEROZOITE.
CC -!- SIMILARITY: STRONG TO AMA-1/PF83 OF P.FALCIPARUM, P.CHABAUDI, AND
CC P.FRAGILE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M58317; AAA63444.1; -
DR EMBL; M61097; AAA29728.1; -
DR InterPro; IPR003298; Apmem_Ag1.
DR Pfam; PF02430; AMA-1; 1.
KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
FT SIGNAL 1 13 POTENTIAL.
FT CHAIN 14 563 MEROZOITE RECEPTOR PK66.
FT DOMAIN 14 487 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 488 508 POTENTIAL.
FT DOMAIN 509 563 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 36 36 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 107 107 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 176 176 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 228 228 N -> K (IN STRAIN LINE W1).
FT SEQUENCE 563 AA; 64680 MW; 9EF0620264AA5CCF CRC64;

Query Match 54.9%; Score 56; DB 1; Length 563;
Best Local Similarity 60.0%; Pred. No. 0.17;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 EKYDKMDEPOHYGKS 18
Db 514 DRYDKMQAEAYGKT 528
:|||||: :|||:

RESULT 7
ID AMA1_PLAFR STANDARD; PRT; 562 AA.
AC P22622; 1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
GN AMA-1 OR AG352.
OS Plasmodium fragile.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
ON NCBI_TaxID=5857;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90205978; PubMed=2181309;
RA Peterson M.G., Nguyen-Dinh P., Marshall V.M., Elliott J.F.,
RT Collins W.E., Anders R.F., Kemp D.J.;
RL "Apical membrane antigen of Plasmodium fragile."
RM Mol. Biochem. Parasitol. 39:279-284(1990).
CC -!- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FALCIPARUM, AND
CC TO PK66 FROM P.KNOWLES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M29898; AAA29474.1; -
DR PIR; A44944; A44944.
DR InterPro; IPR003298; Apmem_Ag1.
DR Pfam; PF02430; AMA-1; 1.
KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.

FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 562 APICAL MEMBRANE ANTIGEN 1.
FT DOMAIN 22 484 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 485 507 POTENTIAL.
FT DOMAIN 508 562 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 176 176 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 226 226 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 562 AA; 64488 MW; 9EAB72D437EA7164 CRC64;

Query Match 48.0%; Score 49; DB 1; Length 562;
Best Local Similarity 57.1%; Pred. No. 2.3;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 EKYDKMDEPOHYGK 17
Db 514 DRYDKMQADGYGK 527
:|||||: :|||:

RESULT 8
ID DLTB_LACCA STANDARD; PRT; 405 AA.
AC P35855;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DLTB PROTEIN (BASIC MEMBRANE PROTEIN) (BMP).
GN DLTB.
OS Lactobacillus casei.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1582;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 7469;
RC MEDLINE=97302479; PubMed=9158726;
RA Neuhaus F.C., Heaton M.P., Debabov D.V., Zhang Q.;
RT "The dlt operon in the biosynthesis of D-alanyl-lipoteichoic acid in
RL Lactobacillus casei."
RN Microb. Drug Resist. 2:77-84(1996).
[2]
RP SEQUENCE OF 1-188 FROM N.A.
RX STRAIN=ATCC 7469;
RC MEDLINE=92325063; PubMed=1385594;
RA Heaton M.P., Neuhaus F.C.;
RT "Biosynthesis of D-alanyl-lipoteichoic acid: cloning, nucleotide
sequence, and expression of the Lactobacillus casei gene for the D-
alanine-activating enzyme."
RL J. Bacteriol. 174:4707-4717(1992).
CC -!- FUNCTION: COULD BE INVOLVED IN THE TRANSPORT OF ACTIVATED
CC D-ALANINE THROUGH THE MEMBRANE.
CC -!- PATHWAY: BIOSYNTHESIS OF D-ALANYL-LIPOTEICHOIC ACID.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U43894; AAB17658.1; -
DR PIR; M83993; AAA25235.1; -
DR PIR; B42983; B42983.
DR Transmembrane; Transport.
KW TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 55 75 POTENTIAL.
FT TRANSMEM 90 110 POTENTIAL.

FT TRANSMEM 116 136 POTENTIAL.
FT TRANSMEM 144 164 POTENTIAL.
FT TRANSMEM 193 213 POTENTIAL.
FT TRANSMEM 225 245 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 317 337 POTENTIAL.
FT TRANSMEM 339 359 POTENTIAL.
FT TRANSMEM 379 399 POTENTIAL.
SQ SEQUENCE 405 AA; 47758 MW; 78A05BC980936F63 CRC64;

Query Match 41.2%; Score 42; DB 1; Length 405;
Best Local Similarity 41.7%; Pred. No. 21;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 YDKMDPEQHYGK 17
DB 174 YDRVDPPEHYAQ 185

RESULT 9
FASD_ECOLI STANDARD; PRT; 835 AA.
AC P46000;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE OUTER MEMBRANE USHER PROTEIN FASD PRECURSOR.
GN FASD.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94148769; PubMed=7906265;
RA Schifferli D.M., Alrutz M.A.;
RT "Permissive linker insertion sites in the outer membrane protein of
RT 987P fimbriae of Escherichia coli."
RT J. Bacteriol. 176:1099-1110(1994).
CC -!- FUNCTION: INVOLVED IN THE EXPORT AND ASSEMBLY OF THE 987P
CC FIMBRIAE SUBUNITS ACROSS THE OUTER MEMBRANE.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE FIMBRIAL EXPORT USHER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; L22659; AAA21827.1; -.
DR EMBL; U50547; AAB02687.1; -.
DR InterPro: IPR000015; Fimb_usher.
DR Pfam: PF00577; Usher; 1
DR PROSITE: PS01151; FIMBRIAL_USHER; 1.
KW Outer membrane; Transmembrane; Fimbria; Transport; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 835 OUTER MEMBRANE USHER PROTEIN FASD.
FT DISULFID 810 834 POTENTIAL.
SQ SEQUENCE 835 AA; 92354 MW; 2FCE6E2305274E1 CRC64;

Query Match 41.2%; Score 42; DB 1; Length 835;
Best Local Similarity 57.1%; Pred. No. 47;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 AEKYDKMDPEQHYG 16
DB 425 ANKYDKIHSGQSYG 438

RESULT 10
VATD_SCHPO STANDARD; PRT; 285 AA.
AC O59823;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE VACUOLAR ATP SYNTHASE SUBUNIT D (EC 3.6.1.34) (V-ATPASE D SUBUNIT)
DE (VACUOLAR PROTON PUMP D SUBUNIT).
GN SPCC965.03.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Wood V., Rajandream M.A., Barrell B.G., Rieger M.;
RL Submitted (May-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SUBUNIT OF THE PERIPHERAL V1 COMPLEX OF VACUOLAR ATPASE.
CC V-ATPASE IS RESPONSIBLE FOR ACIDIFYING A VARIETY OF INTRACELLULAR
CC COMPARTMENTS IN EUKARYOTIC CELLS, THUS PROVIDING MOST OF THE
CC ENERGY REQUIRED FOR TRANSPORT PROCESSES IN THE VACUOLAR SYSTEM.
CC -!- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A
CC PERIPHERAL CATALYTIC V1 COMPLEX (COMPONENTS A TO H) ATTACHED TO
CC AN INTEGRAL MEMBRANE V0 PROTON PORE COMPLEX (COMPONENTS: A, C, C',
CC C'', AND D) (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE V-ATPASE D SUBUNIT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AL023590; CAA19063.1; -.
DR InterPro: IPR002699; ATP-synt_D.
DR Pfam: PF01813; ATP-synt_D; 1.
DR ProDom: PD004122; ATP-synt_D; 1.
KW Hydrolase; ATP synthesis; Hydrogen ion transport.
SQ SEQUENCE 285 AA; 32219 MW; 5B95DF17D1887DD9 CRC64;

Query Match 40.2%; Score 41; DB 1; Length 285;
Best Local Similarity 47.1%; Pred. No. 21;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 NAEKDKMDPEQHYGKS 18
DB 213 NAEKADSVTKKEHQGS 229

RESULT 11
AMPM_CHLPN STANDARD; PRT; 291 AA.
AC Q9Z6Q0; Q9JQD4; Q9K1X1;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE METHIONINE AMINOPEPTIDASE (EC 3.4.11.18) (MAP) (PEPTIDASE M).
GN MAP OR CPN1009 OR CP0844.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydiaophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;

RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RL Nat. Genet. 21:385-389(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heideberg J.F.,
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
 RT pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shiba T., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CWL029 from USA.";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -1- FUNCTION: REMOVES THE AMINO-TERMINAL METHIONINE FROM NASCENT
 CC PROTEINS (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: L-METHIONYLPEPTIDE + H(2)O = L-METHIONINE +
 CC PEPTIDE.
 CC -1- COFACTOR: COBALT (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M24A; ALSO KNOWN AS THE
 CC MAP FAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE001681; AAD19146.1; -
 DR EMBL; AE002244; AAF38634.1; ALT_INT.
 DR EMBL; AP002548; BAA99216.1; -
 DR MEROPS; M24.001; -
 DR TIGR; CP0844; -
 DR InterPro; IPR002467; MAP_1.
 DR InterPro; IPR001714; Methamino_Ptase.
 DR InterPro; IPR000994; Peptidase_M24.
 DR Pfam; PF00557; Peptidase_M24; 1.
 DR PRINTS; PR00599; MAPEPTIDASE.
 DR PROSITE; PS00680; MAP_1; 1.
 DR Hydrolase; AminoPeptidase; Cobalt; Complete proteome.
 FT METAL 135 135 COBALT (BY SIMILARITY).
 FT METAL 146 146 COBALT (BY SIMILARITY).
 FT METAL 209 209 COBALT (BY SIMILARITY).
 FT METAL 241 241 COBALT (BY SIMILARITY).
 FT METAL 274 274 COBALT (BY SIMILARITY).
 SQ SEQUENCE 291 AA; 32596 MW; 094F6D6E17F23DE9 CRC64;

Query Match 40.2%; Score 41; DB 1; Length 291;
 Best Local Similarity 53.8%; Pred. No. 22;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 4 EKDKMDEPOHYG 16
 :||| : |||
 Db 88 KKYDAIAAPHYG 100

RESULT 12
 SCS7_YEAST STANDARD; PRT; 384 AA.
 ID SCS7_YEAST

AC Q03529;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE INOSITOLPHOSPHORYLCERAMIDE-B C-26 HYDROXYLASE (EC 1.-.-.-) (IPC-B
 DE HYDROXYLASE).
 GN SCS7 OR FAH1 OR YMR272C OR YM8156.14C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Lye G., Churcher C.M., Barrell B.G., Rajandream M.A.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=98220296; PubMed=9559540;
 RA Dunn T.M., Haak D., Monaghan E., Beeler T.J.;
 RT "Synthesis of monohydroxylated inositolphosphorylceramide (IPC-C) in
 RT Saccharomyces cerevisiae requires Scs7p, a protein with both a
 RT cytochrome b5-like domain and a hydroxylase/desaturase domain.";
 RL Yeast 14:311-321(1998).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=98019193; PubMed=9353282;
 RA Mitchell A.G., Martin C.E.;
 RT "Fah1p, a Saccharomyces cerevisiae cytochrome b5 fusion protein, and
 RT its Arabidopsis thaliana homolog that lacks the cytochrome b5 domain
 RT both function in the alpha-hydroxylation of sphingolipid-associated
 RT very long chain fatty acids.";
 RL J. Biol. Chem. 272:28281-28288(1997).
 CC -1- FUNCTION: INVOLVED IN THE ALPHA-HYDROXYLATION OF SPHINGOLIPID-
 CC ASSOCIATED VERY LONG CHAIN FATTY ACIDS. HYDROXYLATES THE C26-FATTY
 CC ACID OF INOSITOLPHOSPHORYLCERAMIDE-B (IPC-B) TO FORM IPC-C.
 CC -1- COFACTOR: IRON (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE
 CC AND/OR BE INVOLVED IN METAL ION BINDING.
 CC -1- SIMILARITY: CONTAINS A CYTOCHROME B5 FAMILY, HEME-BINDING DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z49260; CAA89255.1; -
 DR HSSP; P04166; 1B5M.
 DR SGD; S0004885; SCS7.
 DR InterPro; IPR001199; Cyt_B5.
 DR InterPro; IPR001541; Sterol_desat.
 DR Pfam; PF00173; heme_1; 1.
 DR PROSITE; PS00191; CYTOCHROME_B5_1; 1.
 DR PROSITE; PS00255; CYTOCHROME_B5_2; 1.
 KW Oxidoreductase; Fatty acid biosynthesis; Iron; Electron transport;
 KW Transmembrane; Endoplasmic reticulum; Heme.
 FT DOMAIN 1 90 HEME-BINDING.
 FT TRANSMEM 197 217 POTENTIAL.
 FT TRANSMEM 223 243 POTENTIAL.
 FT TRANSMEM 298 318 POTENTIAL.
 FT TRANSMEM 353 373 POTENTIAL.
 FT BINDING 45 45 HEME LIGAND (BY SIMILARITY).
 FT BINDING 70 70 HEME LIGAND (BY SIMILARITY).
 SQ SEQUENCE 384 AA; 44881 MW; DF4BA5F2E0EA2218 CRC64;

Query Match 40.2%; Score 41; DB 1; Length 384;
 Best Local Similarity 54.5%; Pred. No. 29;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 7 DKMDEPOHYCK 17
I:: I:IIII
Db 170 DQIHRPHYCK 180

RESULT 13
AMAL_PLACH STANDARD; PRT; 558 AA.
AC P16445;
DT 01-AUG-1990 (Rel. 15, Last Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
GN AMA-1.
OS Plasmodium chabaudi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5825;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DS;
RX MEDLINE=90114335; PubMed=2608101;
RA Marshall V.M., Peterson M.G., Lew A.M., Kemp D.J.;
RT "Structure of the apical membrane antigen 1 (AMA-1) of Plasmodium
chabaudi.";
RL Mol. Biochem. Parasitol. 37:281-284(1989).
RN [2]
RP SEQUENCE OF 38-377 FROM N.A.
RC STRAIN=DK;
RX MEDLINE=9633375; PubMed=8757869;
RA Crewe P.E., Matthew M.L., Flegg R.H., Anders R.F.;
RT "Protective immune responses to apical membrane antigen 1 of
Plasmodium chabaudi involve recognition of strain-specific
epitopes.";
RL Infect. Immun. 64:3310-3317(1996).
CC -1- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC -1- SUBCELLULAR LOCATION: TYPE 1 MEMBRANE PROTEIN.
CC -1- SIMILARITY: STRONG TO AMA-1/PPF83 FROM P.FALCIPARUM, P.FRAGILE AND
TO PK66 FROM P.KNOWLESII.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M25248; AAA90929.1; -;
CC EMBL; A08270; CAA00765.1; -;
CC EMBL; U49745; AAB36511.1; -;
CC PIR; A44964; A44964.
CC InterPro: IPR003298; Apmem_Ag1.
CC Pfam; PF02430; AMA-1; 1.
CC Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
CC SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 558 APICAL MEMBRANE ANTIGEN 1.
FT DOMAIN 22 480 EXTRACELLULAR (POTENTIAL).
FT TRANSEM 481 503 POTENTIAL.
FT DOMAIN 504 558 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 20 20 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 249 249 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 412 412 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 558 AA; 63973 MW; 9773F3E6A439A372 CRC64;

Query Match 40.2%; Score 41; DB 1; Length 558;
Best Local Similarity 46.7%; Pred. No. 44;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 EKYDKMDEPOHYCKS 18
I:: I:IIII
Db 510 DDYDKMGQADTYGKA 524

RESULT 14
GSPD_ERWCH STANDARD; PRT; 712 AA.
AC P31700;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE GENERAL SECRETION PATHWAY PROTEIN D PRECURSOR (PECTIC ENZYMES
SECRETION PROTEIN OUTD).
GN OUTD.
OS Erwinia chrysanthemi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX NCBI_TaxID=556;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EC16;
RX MEDLINE=93054355; PubMed=1429461;
RA Lindeberg M., Collmer A.;
RT "Analysis of eight out genes in a cluster required for pectic enzyme
secretion by Erwinia chrysanthemi: sequence comparison with secretion
genes from other gram-negative bacteria.";
RL J. Bacteriol. 174:7385-7397(1992).
CC -1- FUNCTION: INVOLVED IN A GENERAL SECRETION PATHWAY (GSP) FOR THE
EXPORT OF PROTEINS. REQUIRED FOR THE TRANSLOCATION OF THE
MULTIPLE PECTIC ENZYMES.
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE EXED/OUTD/PULD/XPSD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L02214; AAA24831.1; -;
CC InterPro: IPR001775; Bac.GSPD.
CC InterPro: IPR000016; Bac.GSPprotein.
CC Pfam; PF00263; Bac.GSPproteins; 1.
CC PRINTS; PR00811; BCTERIALGSPD.
CC PROSITE; PS00875; T2SP.D; 1.
CC TRANSPOS; Outer membrane; Signal.
CC SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 712 GENERAL SECRETION PATHWAY PROTEIN D.
FT DOMAIN 288 353 GLY/SER-RICH.
SQ SEQUENCE 712 AA; 76478 MW; 8A065D9ADA24888 CRC64;
Query Match 40.2%; Score 41; DB 1; Length 712;
Best Local Similarity 50.0%; Pred. No. 57;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 NAEKYDKMDEPOHY 15
I:: I:IIII
Db 65 SVRSYDMDEGQY 78

RESULT 15
GCSP_SCHPO STANDARD; PRT; 1017 AA.
AC Q09785;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PUTATIVE GLYCINE DEHYDROGENASE [DECARBOXYLATING], MITOCHONDRIAL

```

DE PRECURSOR (EC 1.4.4.2) (GLYCINE DECARBOXYLASE) (GLYCINE CLEAVAGE
DE SYSTEM P-PROTEIN).
GN SPAC13G6.06C.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Odell C., Bowman S., Bartell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THE GLYCINE CLEAVAGE SYSTEM CATALYSES THE DEGRADATION OF
CC GLYCINE. THE P PROTEIN BINDS THE ALPHA-AMINO GROUP OF GLYCINE
CC THROUGH ITS PYRIDOXAL PHOSPHATE COFACTOR; CO(2) IS RELEASED AND
CC THE REMAINING METHYLAMINE MOIETY IS THEN TRANSFERRED TO THE
CC LIPOAMIDE COFACTOR OF THE H PROTEIN.
CC -!- CATALYTIC ACTIVITY: GLYCINE + LIPOYLPROTEIN = S-AMINOMETHYL-
CC DIHYDROLIPOYLPROTEIN + CO(2).
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL (BY SIMILARITY).
CC -!- SIMILARITY: TO OTHER GLYCINE CLEAVAGE SYSTEM P-PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z54308; CAA91099.1; -.
DR InterPro; IPR003437; GDC-P.
DR Pfam; PF02347; GDC-P; 1.
KW Hypothetical protein; Oxidoreductase; Mitochondrion;
KW Pyridoxal phosphate; Transit peptide.
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN ? 1017 PUTATIVE GLYCINE DEHYDROGENASE
FT [DECARBOXYLATING].
FT BINDING 769 769 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 1017 AA; 112537 MW; 003CE51DCF104743 CRC64;

```

```

Query Match 40.2%; Score 41; DB 1; Length 1017;
Best Local Similarity 42.1%; Pred. NO. 84;
Matches 8; Conservative 5; Mismatches 4; Indels 2; Gaps 1;
QY 2 NAEKYDKMD--EPQHYGKS 18
Db 39 NSQKFPALDTFEPRHIGPS 57

```

Search completed: January 29, 2002, 11:13:46
Job time: 822 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:12:13 ; Search time 285.36 Seconds
(without alignments)
9.227 Million cell updates/sec

Title: US-09-763-397A-17

Perfect score: 102

Sequence: 1 GNAEKYDKMDEPQHYGKS 18

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	102	100.0	592	5 Q25749	Q25749 plasmodium
2	102	100.0	604	5 Q25746	Q25746 plasmodium
3	102	100.0	622	5 Q9GZ05	Q9GZ05 plasmodium
4	99	97.1	604	5 Q25747	Q25747 plasmodium
5	99	97.1	605	5 Q9GV87	Q9GV87 plasmodium
6	99	97.1	620	5 Q00784	Q00784 plasmodium
7	99	97.1	621	5 Q9TY48	Q9TY48 plasmodium
8	93	91.2	605	5 Q9NFW9	Q9NFW9 plasmodium
9	93	91.2	622	5 Q94661	Q94661 plasmodium
10	90	88.2	592	5 Q25745	Q25745 plasmodium
11	72	70.6	437	5 Q9BHR9	Q9BHR9 plasmodium
12	72	70.6	437	5 Q9BHR8	Q9BHR8 plasmodium
13	72	70.6	437	5 Q9BHR7	Q9BHR7 plasmodium
14	72	70.6	437	5 Q9BHR6	Q9BHR6 plasmodium
15	72	70.6	437	5 Q9BHR5	Q9BHR5 plasmodium
16	72	70.6	437	5 Q9BHR4	Q9BHR4 plasmodium
17	72	70.6	437	5 Q9BHR3	Q9BHR3 plasmodium
18	72	70.6	437	5 Q9BHR0	Q9BHR0 plasmodium
19	72	70.6	437	5 Q9BHQ9	Q9BHQ9 plasmodium

20	72	70.6	437	5 Q9BHQ8	Q9BHQ8 plasmodium
21	72	70.6	437	5 Q9BHQ7	Q9BHQ7 plasmodium
22	72	70.6	437	5 Q9BHQ3	Q9BHQ3 plasmodium
23	72	70.6	437	5 Q9BHQ1	Q9BHQ1 plasmodium
24	72	70.6	437	5 Q9BHQ0	Q9BHQ0 plasmodium
25	72	70.6	437	5 Q9BHP9	Q9BHP9 plasmodium
26	72	70.6	437	5 Q9BHP7	Q9BHP7 plasmodium
27	72	70.6	437	5 Q9BHP4	Q9BHP4 plasmodium
28	72	70.6	437	5 Q9BHP2	Q9BHP2 plasmodium
29	72	70.6	437	5 Q9BHP1	Q9BHP1 plasmodium
30	72	70.6	437	5 Q9BHN9	Q9BHN9 plasmodium
31	72	70.6	437	5 Q9BHN6	Q9BHN6 plasmodium
32	72	70.6	437	5 Q9BHN5	Q9BHN5 plasmodium
33	72	70.6	437	5 Q9BHN4	Q9BHN4 plasmodium
34	72	70.6	437	5 Q9BHN3	Q9BHN3 plasmodium
35	72	70.6	437	5 Q9BHN2	Q9BHN2 plasmodium
36	72	70.6	437	5 Q9BHN6	Q9BHN6 plasmodium
37	72	70.6	437	5 Q9BHB4	Q9BHB4 plasmodium
38	72	70.6	437	5 Q9BHB3	Q9BHB3 plasmodium
39	69	67.6	437	5 Q9BHS1	Q9BHS1 plasmodium
40	69	67.6	437	5 Q9BHS0	Q9BHS0 plasmodium
41	69	67.6	437	5 Q9BHR2	Q9BHR2 plasmodium
42	69	67.6	437	5 Q9BHR1	Q9BHR1 plasmodium
43	69	67.6	437	5 Q9BHQ6	Q9BHQ6 plasmodium
44	69	67.6	437	5 Q9BHQ5	Q9BHQ5 plasmodium
45	69	67.6	437	5 Q9BHQ4	Q9BHQ4 plasmodium

ALIGNMENTS

RESULT 1

ID Q25749	PRELIMINARY;	PRT; 592 AA.
AC Q25749;		
DT 01-NOV-1996	(TREMBLrel. 01, Created)	
DT 01-NOV-1996	(TREMBLrel. 01, Last sequence update)	
DT 01-JUN-2001	(TREMBLrel. 17, Last annotation update)	
DE APICAL MEMBRANE ANTIGEN-1 (FRAGMENT).		
GN AMA-1		
OS Plasmodium falciparum (isolate K1 / Thailand).		
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX NCBI_TaxID=5839;		
RN [1]		
RP SEQUENCE FROM N.A.		
RC STRAIN=K1;		
RX MEDLINE=96379227; PubMed=8784778;		
RA Marshall V.M., Zhang L.X., Anders R.F., Coppel R.L.;		
RT "Diversity of the vaccine candidate AMA-1 of Plasmodium falciparum.";		
RL Mol. Biochem. Parasitol. 77:109-113(1996).		
DR EMBL: U33279; AAC47109.1; "		
DR InterPro: IPR003298; Apmem_Agl.		
DR Pfam; PF02430; AMA_1; 1.		
FT NON_TER	1	
FT NON_TER	592	592
SQ SEQUENCE	592 AA; 68500 MW; F787C3F5306FE877 CRC64;	

Query Match 100.0%; Score 102; DB 5; Length 592;
Best Local Similarity 100.0%; Pred. No. 7.5e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GNAEKYDKMDEPQHYGKS 18

Db 559 GNAEKYDKMDEPQHYGKS 576

RESULT 2

ID Q25746	PRELIMINARY;	PRT; 604 AA.
AC Q25746;		
DT 01-NOV-1996	(TREMBLrel. 01, Created)	
DT 01-NOV-1996	(TREMBLrel. 01, Last sequence update)	

DT 01-JUN-2001.(TREMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN-1 (FRAGMENT).
GN AMA-1.
OS Plasmodium falciparum (isolate vl).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN NCBI_TaxID=5847;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V1;
RX MEDLINE=96379227; PubMed=8784778;
RA Marshall V.M., Zhang L.X., Anders R.F., Coppel R.L.;
RT "Diversity of the vaccine candidate AMA-1 of Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 77:109-113(1996).
DR EMBL: U33276; AAC47106.1; -;
DR InterPro: IPR003298; Apmem_Ag1.
DR Pfam: PF02430; AMA-1; 1.
FT NON_TER 604 604
SQ SEQUENCE 604 AA; 69892 MW; CB83EE928376BFFB CRC64;

Query Match 100.0%; Score 102; DB 5; Length 604;
Best Local Similarity 100.0%; Pred. No. 7.7e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPOHYGKS 18
|||||
Db 571 GNAEKYDKMDEPOHYGKS 588

RESULT 3
Q9GZ05 PRELIMINARY; PRT; 622 AA.
AC Q9GZ05;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1.
GN AMA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN NCBI_TaxID=5833;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FCCL/HN;
RA Shan Z.X., Yu X.B., Li X.R., Ma C.L., Fang J.M.;
RT "Cloning and structure of apical membrane antigen 1 (AMA-1) gene of Plasmodium falciparum isolate FCCL/HN.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF277003; AAG01325.1; -;
DR InterPro: IPR003298; Apmem_Ag1.
DR Pfam: PF02430; AMA-1; 1.
DR PRINTS: PR01361; MEROZOITESA.
SQ SEQUENCE 622 AA; 72045 MW; 69045E66776C3BBB CRC64;

Query Match 100.0%; Score 102; DB 5; Length 622;
Best Local Similarity 100.0%; Pred. No. 7.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPOHYGKS 18
|||||
Db 571 GNAEKYDKMDEPOHYGKS 588

RESULT 4
Q25747 PRELIMINARY; PRT; 604 AA.
AC Q25747;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN-1 (FRAGMENT).
GN AMA-1.

OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN NCBI_TaxID=5833;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB3;
RX MEDLINE=96379227; PubMed=8784778;
RA Marshall V.M., Zhang L.X., Anders R.F., Coppel R.L.;
RT "Diversity of the vaccine candidate AMA-1 of Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 77:109-113(1996).
DR EMBL: U33277; AAC47107.1; -;
DR InterPro: IPR003298; Apmem_Ag1.
DR Pfam: PF02430; AMA-1; 1.
FT NON_TER 604 604
SQ SEQUENCE 604 AA; 69886 MW; 791620C3B24D342B CRC64;

Query Match 97.1%; Score 99; DB 5; Length 604;
Best Local Similarity 94.4%; Pred. No. 2.3e-07;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPOHYGKS 18
|||||
Db 571 GNAEKYDKMDEPOHYGKS 588

RESULT 5
Q9GV87 PRELIMINARY; PRT; 605 AA.
AC Q9GV87;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN NCBI_TaxID=5833;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVO;
RA Kocken C.H.M., Dubbeld M.A., Herrera S., Thomas A.W.;
RT "Diversity of Apical Membrane Antigen 1 sequence in Plasmodium falciparum FVO strain.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ277646; CAC05390.1; -;
DR InterPro: IPR003298; Apmem_Ag1.
DR Pfam: PF02430; AMA-1; 1.
DR PRINTS: PR01361; MEROZOITESA.
FT NON_TER 605 605
SQ SEQUENCE 605 AA; 69897 MW; F6D89F02F7690C19 CRC64;

Query Match 97.1%; Score 99; DB 5; Length 605;
Best Local Similarity 94.4%; Pred. No. 2.3e-07;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPOHYGKS 18
|||||
Db 563 GNAEKYDKMDEPOHYGKS 580

RESULT 6
O00784 PRELIMINARY; PRT; 620 AA.
ID O00784;
AC O00784;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN-1 (FRAGMENT).
GN AMA-1.
OS Plasmodium falciparum (isolate kfl916).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=57269;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN-KF1916;
 RA Eisen D., Billman-Jacobe H., Marshall V.M., Coppel R.L.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN-KF1916;
 RA Eisen D.P., Marshall V.M., Billman-Jacobe H., Coppel R.L.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U0276; AAB50405.1; -;
 DR EMBL; AF061332; AAC15773.1; -;
 DR InterPro; IPR003298; Apmem_Agl.
 DR Pfam; PF02430; AMA-1; 1.
 FT NON_TER 520 620
 SQ SEQUENCE 620 AA; 71/721 MW; 9194F06A0D138677 CRC64;

Query Match 97.1%; Score 99; DB 5; Length 620;
 Best Local Similarity 94.4%; Pred. No. 2.3e-07;
 Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GNAEKYDKMDPQHYGKS 18
 |||||
 Db 571 GNAEKYDKMDPQHYGKS 588

RESULT 7
 ID Q9TY48 PRELIMINARY; PRT; 621 AA.
 AC Q9TY48;
 DT 01-MAY-2000 (TrEMBLrel. 13; Created)
 DT 01-MAY-2000 (TrEMBLrel. 13; Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17; Last annotation update)
 DE APICAL MEMBRANE ANTIGEN-1 (FRAGMENT).
 GN AMA-1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN-FVO;
 RA Aguilar J.C., Hoffman S.L.;
 RT "The immunogenicity and protective efficacy of Plasmodium falciparum AMA-1 DNA vaccines";
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U84348; AAD03790.1; -;
 DR InterPro; IPR003298; Apmem_Agl.
 DR Pfam; PF02430; AMA-1; 1.
 FT NON_TER 521 621
 SQ SEQUENCE 621 AA; 71921 MW; 6E2A28C5142D8A7D CRC64;

Query Match 97.1%; Score 99; DB 5; Length 621;
 Best Local Similarity 94.4%; Pred. No. 2.3e-07;
 Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GNAEKYDKMDPQHYGKS 18
 |||||
 Db 570 GNAEKYDKMDPQHYGKS 587

RESULT 8
 ID Q9NFW9 PRELIMINARY; PRT; 605 AA.
 AC Q9NFW9;
 DT 01-OCT-2000 (TrEMBLrel. 15; Created)
 DT 01-OCT-2000 (TrEMBLrel. 15; Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17; Last annotation update)
 DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
 GN 83/AMA-1.

OS Plasmodium reichenowi.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5854;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP Kocken C.H.M., Dubbeld M.A., Thomas A.W.;
 RT "Apical membrane antigen 1 of Plasmodium reichenowi: molecular characterisation, comparison with Plasmodium falciparum AMA-1 and antibody-mediated inhibition";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ252087; CAB66387.1; -;
 DR InterPro; IPR003298; Apmem_Agl.
 DR Pfam; PF02430; AMA-1; 1.
 DR PRINTS; PR01361; MEROZOITESA.
 FT NON_TER 1 605
 FT NON_TER 605 605
 SQ SEQUENCE 605 AA; 70072 MW; E4959A99D5C880C6 CRC64;

Query Match 91.2%; Score 93; DB 5; Length 605;
 Best Local Similarity 94.1%; Pred. No. 2e-06;
 Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NAEKYDKMDPQHYGKS 18
 |||||
 Db 564 NAEKYDKMDPQHYGKS 580

RESULT 9
 ID Q94661 PRELIMINARY; PRT; 622 AA.
 AC Q94661; Q25744;
 DT 01-FEB-1997 (TrEMBLrel. 02; Created)
 DT 01-FEB-1997 (TrEMBLrel. 02; Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17; Last annotation update)
 DE APICAL MEMBRANE ANTIGEN 1 (APICAL MEMBRANE ANTIGEN-1).
 GN AMA-1.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=3D7;
 RA Tine J.A., Lanar D.E., Smith D.M., Wellde B.T., Schultheiss P., Ware L.A., Kaufman E.B., Wirtz R.A., de Taisne C., Hui G.S.N., Chang S.P., Church P., Hollingdale M.R., Kaslow D.C., Hoffman S., Guito K.P., Ballou W.R., Sadoff J.C., Paoletti E.;
 RA Infect. Immun. 64:0-0(0).
 RL [2]
 RP SEQUENCE OF 9-604 FROM N.A.
 RC STRAIN=3D7;
 RX MEDLINE=96379227; PubMed=8784778;
 RA Marshall V.M., Zhang L.X., Anders R.F., Coppel R.L.;
 RT "Diversity of the vaccine candidate AMA-1 of Plasmodium falciparum.";
 RL Mol. Biochem. Parasitol. 77:109-113(1996).
 DR EMBL; U65407; AAB36701.1; -;
 DR EMBL; U33274; AAC47104.1; -;
 DR InterPro; IPR003298; Apmem_Agl.
 DR Pfam; PF02430; AMA-1; 1.
 DR PRINTS; PR01361; MEROZOITESA.
 SQ SEQUENCE 622 AA; 72041 MW; 60FE442074C38E94 CRC64;

Query Match 91.2%; Score 93; DB 5; Length 622;
 Best Local Similarity 94.4%; Pred. No. 2e-06;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GNAEKYDKMDPQHYGKS 18
 |||||
 Db 571 GNAEKYDKMDPQHYGKS 588

RESULT 10

Q25745	PRELIMINARY; PRT; 592 AA.
ID Q25745	
AC Q25745;	
DT 01-NOV-1996 (TrEMBLrel. 01, Created)	
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)	
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)	
DE APICAL MEMBRANE ANTIGEN-1 (FRAGMENT).	
GN AMA-1.	
OS Plasmodium falciparum.	
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
OX NCBI_TaxID=5833;	
LN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=CMPI;	
RX MEDLINE=96147889; PubMed=8556798;	
RA Zhang L.X., Zhan B., Wang J., Feng X.;	
RL Chung Kuo Chi Sheng Chung Hsueh Yu Chi Sheng Chung 13:203-208(1995).	
RN [2]	
RP SEQUENCE FROM N.A.	
RC STRAIN=CMPI;	
RX MEDLINE=96379227; PubMed=8784778;	
RA Marshall V.M., Zhang L.X., Anders R.F., Coppel R.L.;	
RT "Diversity of the vaccine candidate AMA-1 of Plasmodium falciparum.";	
RL Mol. Biochem. Parasitol. 77:109-113(1996).	
DR EMBL; U33275; AAC47105.1; -	
DR InterPro; IPR003298; Apmem_Agl.	
DR Pfam; PF02430; AMA-1; 1.	
FT NON_TER 1	
FT NON_TER 592 592	
FT NON_TER 592 592	
SQ SEQUENCE 592 AA; 68412 MW; 5496A73680E7128C CRC64;	
Query Match 88.2%; Score 90; DB 5; Length 592;	
Best Local Similarity 88.9%; Pred. No. 5.6e-06;	
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	
QY 1 GNAEKYDKMDPEQHYGKS 18	
Db 559 GNAEKYDKMDPQDYGKS 576	
RESULT 11	
Q9BHR9 PRELIMINARY; PRT; 437 AA.	
ID Q9BHR9	
AC Q9BHR9;	
DT 01-JUN-2001 (TrEMBLrel. 17, Created)	
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)	
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)	
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).	
GN AMA1.	
OS Plasmodium falciparum.	
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
OX NCBI_TaxID=5833;	
LN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=NIGERIAN 015;	
RA Polley S.D., Conway D.J.;	
RT "Diversifying selection in domains of Plasmodium falciparum Apical	
RL Membrane Antigen 1 (AMA1).";	
RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.	
DR EMBL; AJ408303; CAC34744.1; -	
DR NON_TER 1	
DR NON_TER 437 437	
DR NON_TER 437 437	
SQ SEQUENCE 437 AA; 50243 MW; 28D9E9E5165A39EE CRC64;	
Query Match 70.6%; Score 72; DB 5; Length 437;	
Best Local Similarity 100.0%; Pred. No. 0.0027;	
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1 GNAEKYDKMDPEQ 13	
Db 425 GNAEKYDKMDPEQ 437	

AC Q9BHR6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NIGERIAN 034A;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical
RT Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ408308; CAC34749.1; -.
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 50217 MW; 68A41FEE41C3271C CRC64;

Query Match 70.6%; Score 72; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPQ 13
|||||
DB 425 GNAEKYDKMDEPQ 437

RESULT 15
Q9BHR5
ID Q9BHR5 PRELIMINARY; PRT; 437 AA.
AC Q9BHR5;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NIGERIAN 034B;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical
RT Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ408309; CAC34750.1; -.
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 50206 MW; EB05396E7D92AC98 CRC64;

Query Match 70.6%; Score 72; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPQ 13
|||||
DB 425 GNAEKYDKMDEPQ 437

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:21:49 ; Search time 310.82 seconds
(without alignments)
4.528 Million cell updates/sec

Title: US-09-763-397A-18

Perfect score: 105

Sequence: 1 DQPKQYEQHLTDYKIKEG 19

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_1101.*
- 1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
 - 2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
 - 3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
 - 4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
 - 5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
 - 6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
 - 7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
 - 8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
 - 9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
 - 10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
 - 11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
 - 12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
 - 13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
 - 14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
 - 15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
 - 16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
 - 17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
 - 18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
 - 19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
 - 20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
 - 21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
 - 22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	100.0	19	AAV70294	Plasmodium falcipa
2	105	100.0	350	AAV70278	Recombinant vaccin
3	105	100.0	622	AAV70278	Rhoptry membrane a
4	105	100.0	622	AAV70278	Plasmodium falcipa
5	105	100.0	622	AAV70278	Plasmodium falcipa
6	87	82.9	563	AAV70278	Plasmodium knowles
7	79	75.2	489	AAV70278	Plasmodium vivax 6
8	72	68.6	561	AAV70278	Rhoptry membrane a
9	53	50.5	360	AAV70278	Human SGP039 phosph
10	53	50.5	551	AAV70278	Mouse serine/threo
11	53	50.5	553	AAV70278	Novel protein kina

12	53	50.5	568	19	AAV75791	Human protein kina
13	53	50.5	568	21	AAV75791	HYAK3-2 polypeptid
14	53	50.5	568	22	AAV75791	Human YAK3b (HYAK3
15	53	50.5	588	19	AAV75790	Human protein kina
16	53	50.5	588	21	AAV75790	HYAK3-2 alpha poly
17	53	50.5	588	21	AAV75792	Amino acid sequenc
18	53	50.5	588	22	AAV75792	Human YAK3a (HYAK3
19	51	48.6	323	18	AAV21900	Cotton fibre struc
20	50.5	48.1	681	21	AAV19283	Amino acid sequenc
21	45	42.9	95	21	AAV19096	Zea mays protein f
22	45	42.9	139	21	AAV19095	Zea mays protein f
23	44	41.9	308	19	AAV59131	Homo sapiens Tub I
24	44	41.9	405	21	AAV58737	Breast and ovarian
25	44	41.9	458	14	AAV41781	Rabbit HBI protein
26	44	41.9	459	19	AAV40338	Human FK506 bindin
27	44	41.9	484	20	AAV28467	Tetratricopeptide
28	44	41.9	484	20	AAV17247	Human protein kina
29	44	41.9	484	22	AAV72677	Human tetratricope
30	44	41.9	484	22	AAV72682	Human tetratricope
31	44	41.9	1865	17	AAV03515	Human DOCK180 prot
32	43	41.0	109	22	AAV82336	S. epidermidis ope
33	43	41.0	132	22	AAV81810	S. epidermidis ope
34	43	41.0	248	21	AAV32739	Eucalyptus grandis
35	43	41.0	295	20	AAV34850	Chlamydia pneumoni
36	43	41.0	436	22	AAV94564	Human protein sequ
37	43	41.0	456	22	AAV82048	S. epidermidis ope
38	43	41.0	1349	20	AAV16434	Multiple drug resi
39	43	41.0	1588	15	AAV46605	Malarial PfEMP3 ep
40	43	41.0	1663	15	AAV46608	Plasmodium falcipa
41	43	41.0	2441	21	AAV18161	Plasmodium falcipa
42	42	40.0	86	22	AAV46484	B. licheniformis b
43	42	40.0	87	22	AAV01304	Human secreted pro
44	42	40.0	89	22	AAV60091	Human transport pr
45	42	40.0	105	22	AAV14030	Peptide #464 encod

ALIGNMENTS

RESULT 1

AAV70294

ID AAV70294 standard; peptide: 19 AA.

XX AC AAV70294;

XX XX

DT 06-JUN-2000 (first entry)

XX XX

DE Plasmodium falciparum AMA-1 antigenic epitope, P602.

XX XX

KW Recombinant protein; CDC/NIAIDVAC-1; multivalent; malaria; vaccine;

KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;

KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;

KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1;

KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;

KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;

KW Pf27; antiparasitic; prevention; anti-CDC/NIAIDVAC-1 antibody.

XX OS Plasmodium falciparum.

XX XX

PN WO200011179-A1.

XX XX

PD 02-MAR-2000.

XX XX

PF 19-AUG-1999; 99WO-US18869.

XX XX

PR 21-AUG-1998; 98US-0097703.

XX XX

PA (NAIM-) NAT INST IMMUNOLOGY.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX XX

PI Lal AA, Shi YP, Hasnain SE;

XX XX

DR WPI; 2000-237654/20.

XX Novel recombinant protein as vaccine for treating malarial infection
 PT comprises antigenic peptides obtained from different stages of
 PT Plasmodium falciparum life cycle -
 XX
 XX Claim 2; Page 17; 52pp; English.
 PS
 XX The present sequence is the antigenic epitope P602, derived from
 CC apical membrane antigen-1 (AMA-1) of the asexual blood stage of
 CC Plasmodium falciparum. It is used in the construction of recombinant
 CC protein CDC/NiMALVAC-1, which is a multivalent, multistage malarial
 CC vaccine. The recombinant protein comprises, melittin signal peptide,
 CC (His)6 tag, T-cell epitope from tetanus toxoid and 21 antigenic epitopes
 CC from circumsporozoite protein (CSP), sporozoite surface protein-2
 CC (SSP-2), liver stage antigen-1 (LSA-1), merozoite surface protein-1
 CC (MSP-1), apical membrane antigen-1 (AMA-1), erythrocyte binding
 CC antigen-175 (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete
 CC specific antigen, Pf27. These epitopes were obtained at different stages
 CC of the life cycle of P. falciparum. CDC/NiMALVAC-1 vaccine has
 CC antiparasitic activity and can be used for treatment and prevention of
 CC malarial infections. Anti-CDC/NiMALVAC-1 antibodies can be used for
 CC detecting P. falciparum in biological samples.
 XX
 XX Sequence 19 AA;

Query Match 100.0%; Score 105; DB 21; Length 19;
 Best Local Similarity 100.0%; Pred. No. 7.9e-09;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQYEQHLTDYKIKEG 19

Db 1 dqpkyeqhltdyekikey 19

RESULT 2

AAY70278

ID AAY70278 standard; Protein; 350 AA.

XX

AC AAY70278;

XX 06-JUN-2000 (first entry)

XX Recombinant vaccine CDC/NiMALVAC-1.

XX Recombinant protein; CDC/NiMALVAC-1; multivalent; malaria; vaccine;
 KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
 KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
 KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
 KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
 KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
 KW Pf27; antiparasitic; prevention; anti-CDC/NiMALVAC-1 antibody;
 KW honey bee.

XX Chimeric - Apis sp.

OS Chimeric - Clostridium tetani.

OS Chimeric - Plasmodium falciparum.

XX Key Location/Qualifiers

FT Peptide 1..22

FT /label= Melittin_signal_peptide

FT /note= "Derived from Honey Bee"

FT Protein 23..350

FT /label= Mature_CDC/NiMALVAC-1

FT /note= "Recombinant multivalent malarial vaccine"

XX WO200011179-A1.

XX 02-MAR-2000.

XX 19-AUG-1999; 99WO-US18869.

XX 21-AUG-1998; 98US-0097703.

XX

PA

PA

XX

PI

XX

DR

DR

XX

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

(NAIM-) NAT INST IMMUNOLOGY.
 (USSH) US DEPT HEALTH & HUMAN SERVICES.

Lal AA, Shi YP, Hasnain SE;

WPI; 2000-237654/20.

N-PSDB; AAZ51336.

Novel recombinant protein as vaccine for treating malarial infection
 comprises antigenic peptides obtained from different stages of
 Plasmodium falciparum life cycle -

Claim 3; Page 43-44; 52pp; English.

The present sequence is that of recombinant protein CDC/NiMALVAC-1,
 which is a multivalent, multistage malarial vaccine. The recombinant
 protein comprises, melittin signal peptide, (His)6 tag, T-cell epitope
 from tetanus toxoid and 21 antigenic epitopes from circumsporozoite
 protein (CSP), sporozoite surface protein-2 (SSP-2), liver stage
 antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1), MSP-2, apical
 membrane antigen-1 (AMA-1), erythrocyte binding antigen-175 (EBA-175),
 rhoptry associated protein-1 (RAP-1) and gamete specific antigen, Pf27.
 These epitopes were obtained at different stages of the life cycle of
 Plasmodium falciparum. CDC/NiMALVAC-1 vaccine has antiparasitic
 activity and can be used for treatment and prevention of malarial
 infections. Anti-CDC/NiMALVAC-1 antibodies can be used for detecting
 P. falciparum in biological samples.

Sequence 350 AA;

Query Match 100.0%; Score 105; DB 21; Length 350;

Best Local Similarity 100.0%; Pred. No. 1.7e-07;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQYEQHLTDYKIKEG 19

Db 264 dqpkyeqhltdyekikey 282

RESULT 3

APP91632

ID AAP91632 standard; protein; 622 AA.

XX

AC AAP91632;

XX 25-JAN-1990 (first entry)

XX Rhoptry membrane antigen-1.

XX Malaria; rhoptry membrane antigen-1; antibodies.

XX Plasmodium falciparum D10.

XX WO8907645-A.

XX 24-AUG-1989.

XX 10-FEB-1989; 89WO-AU00056.

XX 12-FEB-1988; 89AU-0006743.

XX (SARA) SARAMANE PTY LTD.

XX Peterson MG, Crewther PE, Smythe JA, Marshall VM, Silva A;

XX WPI; 1989-263714/36.

XX N-PSDB; AAN90703.

XX Rhoptry membrane antigen of Plasmodium falciparum

XX - used for producing antibodies and in immunisation,

XX diagnostic and treatment methods for malaria.

XX Claim 1; Fig 3; 46pp; English.

CC RMA-1 can generate an immune response to malaria, and antibodies which can inhibit growth of the parasite. RMA-1 initially has mol. wt. 80 kD.

SQ Sequence 622 AA;

Query Match 100.0%; Score 105; DB 10; Length 622;
Best Local Similarity 100.0%; Pred. No. 3e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQVEQHLTDYEKKEG 19
|||||
Db 348 dqpkgyeqhltdyekikeg 366

RESULT 4

AAAR27532
ID AAR27532 standard; Protein; 622 AA.

AC AAR27532;

DT 08-MAR-1993 (first entry)

DE Plasmodium falciparum AMA1 antigen.

KW Recombinant poxvirus; antimalarial vaccine; malaria; immunise;
KW immunogen.

OS Plasmodium falciparum.

PN WO9216616-A.

PD 01-OCT-1992.

PF 19-MAR-1992; 92WO-US02207.

PR 20-MAR-1991; 91US-0672183.

PR 18-MAR-1992; 92US-0852305.

PA (VIRO-) VIROGENETICS CORP.

PI De Taisne C, Paoletti E, Tine JA;

DR WPI; 1992-349203/42.

DR N-PSDB; AAQ29189.

PT Recombinant poxvirus - contg. Plasmodium DNA, useful as
PT antimalarial vaccine

PS Example 4; Fig 5; 74pp; English.

CC This sequence is the Plasmodium falciparum AMA1 antigen.
CC cDNA encoding it was cloned into vaccinia donor plasmids
CC before being inserted into the vaccinia virus to be used in a
CC vaccine to stimulate an antimalarial immunological response, or for
CC in vitro prodn. of gene prods. for use as immunogens. As plasmodium
CC genes are conserved among P. falciparum strains, they are widely
CC effective in a vaccine.

SQ Sequence 622 AA;

Query Match 100.0%; Score 105; DB 13; Length 622;
Best Local Similarity 100.0%; Pred. No. 3e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQVEQHLTDYEKKEG 19
|||||
Db 348 dqpkgyeqhltdyekikeg 366

RESULT 5

AAAR68840

ID AAR68840 standard; Protein; 622 AA.

AC AAR68840;

DT 24-AUG-1995 (first entry)

DE Plasmodium falciparum AMA-1 gene protein.

KW Plasmodium falciparum AMA-1 gene; recombinant poxvirus;
KW multicomponent multistage malarial vaccines; immunogens;
KW malaria diagnosis.

OS Plasmodium falciparum (3D7).

PN WO9428930-A.

PD 22-DEC-1994.

PF 10-JUN-1994; 94WO-US06652.

PR 11-JUN-1993; 93US-0075783.

PR 09-JUN-1994; 94US-0257073.

PA (VIRO-) VIROGENETICS CORP.

PI De Taisne C, Paoletti E, Tine JA;

DR WPI; 1995-036113/05.

DR N-PSDB; AAQ80910.

PT Recombinant poxvirus contg. Plasmodium DNA in non-essential
PT region - useful in vaccines against malaria and for prodn. of
PT Plasmodium immunogens

PS Claim 3; Fig 5; 183pp; English.

CC AAQ80910 encodes AAR68840 the P. falciparum AMA-1 gene product. New
CC recombinant poxviruses containing either the SERA, ABRA, Pfmsp70,
CC AMA-1, Pf25, Pf16, CSP, PfSSP2, LSA-1, LSA-1 repeats, MSA-1,
CC MSA-1 (N-terminal p83 or C-terminal gp42) genes, or a combination
CC of these in non-essential regions of their genomes are claimed.
CC These poxviruses (pref. with a virulence reducing genomic
CC deletion or disruption) can be used as vaccines against malaria
CC and for the prodn. of Plasmodium immunogens. These viruses
CC provide multicomponent, multistage vaccines due to their expression
CC of sporozoite, liver stage, blood stage and sexual stage proteins.

SQ Sequence 622 AA;

Query Match 100.0%; Score 105; DB 16; Length 622;
Best Local Similarity 100.0%; Pred. No. 3e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQVEQHLTDYEKKEG 19
|||||
Db 348 dqpkgyeqhltdyekikeg 366

RESULT 6

AAAR10551

ID AAR10551 standard; Protein; 563 AA.

AC AAR10551;

DT 15-APR-1991 (first entry)

DE Plasmodium knowlesi 66kD merozoite antigen protein.

KW Plasmodium merozoite antigens; malaria vaccine.

XX OS Plasmodium knowlesi.
 XX PN US7483516-A.
 XX PD 15-JAN-1991.
 XX PF 22-FEB-1990; 90US-0483516.
 XX PR 22-FEB-1990; 90US-0483516.
 XX PA (USSH) NAT INST OF HEALTH.
 XX PI Waters AP, McCutchan TF;
 XX DR WPI; 1991-044381/06.
 XX DR N-PSDB; AAQ10418.
 XX PT Plasmodium merozoite antigen proteins and DNA sequences - useful
 PT in vaccine prodn., anti-malarial drug design, and in diagnostics
 XX PS Disclosure; fig 1; 3lpp; English.
 XX CC This protein is encoded by a recombinant DNA molecule comprising
 CC the complete P.knowlesi merozoite antigen gene and e.g. a pUC19
 CC or vaccinia virus vector sequence and regulatory elements. This
 CC allows efficient expression of the antigen on transformation of
 CC host cells. It is useful in an anti-malarial vaccine for con-
 CC ferring immunity against the merozoite form of the parasite and
 CC preventing infection of uninfected red blood cells.
 XX CC See also AAQ10869.
 XX SQ Sequence 563 AA;

Query Match 82.9%; Score 87; DB 12; Length 563;
 Best Local Similarity 78.9%; Pred. No. 0.00011;
 Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 DOPKQYEQHLTDYEKIKEG 19
 |||:||||:|||||:|
 Db 293 dqprqyeeltdyeki qeg 311

RESULT 7
 AAR10935
 ID AAR10935 standard; Protein; 489 AA.
 XX AC AAR10935;
 XX DT 15-APR-1991 (first entry)
 XX DE Plasmodium vivax 66kD merozoite antigen protein partial sequence.
 XX KW Plasmodium merozoite antigens; malaria vaccine.
 XX OS Plasmodium vivax.
 XX PN US7483516-A.
 XX PD 15-JAN-1991.
 XX PF 22-FEB-1990; 90US-0483516.
 XX PR 22-FEB-1990; 90US-0483516.
 XX PA (USSH) NAT INST OF HEALTH.
 XX PI Waters AP, McCutchan TF;
 XX DR WPI; 1991-044381/06.
 XX DR N-PSDB; AAQ10869.

PT Plasmodium merozoite antigen proteins and DNA sequences - useful
 PT in vaccine prodn., anti-malarial drug design, and in diagnostics
 XX PS Disclosure; fig 1; 3lpp; English.
 XX CC This protein is encoded by a recombinant DNA molecule comprising
 CC the complete P.vivax merozoite antigen gene and e.g. a pUC19
 CC or vaccinia virus vector sequence and regulatory elements. This
 CC allows efficient expression of the antigen on transformation of
 CC host cells. It is useful in an anti-malarial vaccine for con-
 CC ferring immunity against the merozoite form of the parasite and
 CC preventing infection of uninfected red blood cells.
 XX CC See also AAQ10418.
 XX SQ Sequence 489 AA;

Query Match 75.2%; Score 79; DB 12; Length 489;
 Best Local Similarity 63.2%; Pred. No. 0.0014;
 Matches 12; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 1 DOPKQYEQHLTDYEKIKEG 19
 |||:||||:|||||:|
 Db 219 dqprqyeeltdyeki qeg 237

RESULT 8
 AAP91588
 ID AAP91588 standard; protein; 561 AA.
 XX AC AAP91588;
 XX DT 25-JAN-1990 (first entry)
 XX DE Rhoptry membrane antigen-1.
 XX KW Malaria; rhoptry membrane antigen-1; antibodies.
 XX OS Plasmodium chabaudi adami DS.
 XX PN WO8907645-A.
 XX PD 24-AUG-1989.
 XX PF 10-FEB-1989; 89WO-AU000056.
 XX PR 12-FEB-1988; 88AU-0006743.
 XX PA (SARA) SARAMANE PTY LTD.
 XX PI Peterson MG, Crewther PE, Smythe JA, Marshall VM, Silva A;
 XX DR WPI; 1989-263714/36.
 XX DR N-PSDB; AAP91632.
 XX KW Rhoptry membrane antigen of Plasmodium falciparum
 PT - used for producing antibodies and in immunisation,
 PT diagnostic and treatment methods for malaria.
 XX PS Disclosure; Fig. 3; 46pp; English.
 XX CC Predicted mol. wt. is 63901 D. This shows great similarity to
 CC P. falciparum RMA-1 protein (sequence in specification-claimed).
 XX SQ Sequence 561 AA;

Query Match 68.6%; Score 72; DB 10; Length 561;
 Best Local Similarity 68.4%; Pred. No. 0.017;
 Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 DOPKQYEQHLTDYEKIKEG 19
 |||:||||:|||||:|

Db 293 dqpkyekhledtakirrg 311

RESULT 9
AAE04841
ID AAE04841 standard; Protein: 360 AA.
XX
AC AAE04841;
XX
XX
DT 10-SEP-2001 (first entry)
XX
XX Human SGP039 phosphatase polypeptide.
DE
XX
XX Human; SGP039 phosphatase polypeptide; phosphatase-related disease;
KW immune-related disorder; ocular disease; organ transplant rejection;
KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
KW metabolic disorder; haematopoietic cancer; mood disorder; cardiac;
KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
KW attention disorder; cognitive disorder; psychotic disorder; cytostatic;
KW neurological disorder; viricide; nootropic; cerebroprotective; therapy;
KW neuroprotective; antibacterial; vulnery; tranquilliser; antiasthmatic;
KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
KW antifungal; migraine; serine phosphatase; STP; PP2C.
XX
XX Homo sapiens.
OS
XX
XX
XX Key Location/Qualifiers
FH Domain 1..301
FT /label= Catalytic_domain
FT 91..344
FT /label= Phosphatase_domain
XX
XX WO200146394-A2.
XX
XX 28-JUN-2001.
XX
XX 21-DEC-2000; 2000WO-US34736.
XX
XX 21-DEC-1999; 99US-0173255.
PR 28-DEC-1999; 99US-0175766.
PR 25-JAN-2000; 2000US-0178078.
PR 31-JAN-2000; 2000US-0179301.
XX
XX (SUGE-) SUGEN INC.
XX
XX Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ;
PI Flanagan P;
XX
XX WPI; 2001-418058/44.
DR N-PSDB; AAD09499.
XX
XX Novel phosphatase polypeptide useful for treating cancers,
PT immune-related diseases and disorders, cardiovascular disease, brain or
PT neuronal-associated diseases and metabolic disorders
XX
PS Claim 7; Fig 2; 186pp; English.
XX
XX The present invention relates to phosphatase polypeptides, nucleotide
CC sequences encoding them, as well as various products and methods useful
CC for the diagnosis and treatment of various phosphatase-related diseases
CC and conditions. Substance that modulates the activity of phosphatase
CC polypeptide is used to treat immune-related diseases and disorders,
CC cardiovascular disease, brain or neuronal-associated diseases and
CC metabolic disorders, including cancers of tissues, cancers of
CC haematopoietic origin, diseases of central and peripheral nervous
CC system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,
CC amyotrophic lateral sclerosis, viral infections, infections caused by
CC prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain,
CC sexual dysfunction, mood disorders, attention disorders, cognition
CC disorders, hypotension, hypertension, psychotic disorders, neurological
CC disorders, dyskinesias and organ transplant rejection. The present
CC amino acid sequence is human SGP039 phosphatase polypeptide. This

CC sequence is classified as serine phosphatase, STP and PP2C.
XX
SQ Sequence 360 AA;
Query Match 50.5%; Score 53; DB 22; Length 360;
Best Local Similarity 62.5%; Pred. No. 5.9;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy 3 PKQYEQHLTDYEKKE 18
I: :||| |||| ||
Db 144 pealkqlqdyekdke 159

RESULT 10
AAE02843
ID AAE02843 standard; Protein: 551 AA.
XX
AC AAE02843;
XX
DT 10-AUG-2001 (first entry)
XX
DE Mouse serine/threonine kinase, YAK3 version #1.
XX
XX Mouse; serine/threonine kinase; YAK3; therapy; neutropenia; cytopenia;
KW anaemia; renal insufficiency; chronic disease; autoimmunity; cancer;
KW polycythaemia; myelosuppression; infertility; testicular cancer;
KW bone loss; osteoporosis; inflammatory disease; rheumatoid arthritis;
KW psoriasis; adult respiratory disease syndrome; dermatitis; asthma;
KW allergy; infection; human immunodeficiency virus; HIV; pain; injury;
KW immunodeficiency disorder; anorexia; Parkinson's disease; hypotension;
KW atherosclerosis; cardiovascular disease; hypertension; ulcer; psychotic;
KW myocardial infarction; neurological disorder; anxiety; schizophrenia;
KW dementia; delirium; dyskinesias; Huntington's disease.
XX
XX Mus musculus.
OS
XX
XX WO200132889-A1.
XX
XX 10-MAY-2001.
XX
XX 06-NOV-2000; 2000WO-US30505.
XX
XX 04-NOV-1999; 99US-0163621.
PR 03-NOV-2000; 2000US-0706385.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX (SMIK) SMITHKLINE BEECHAM PLC.
XX
XX Creasy CL, Hughes SA, Wojchowski DM;
XX
XX WPI; 2001-335832/35.
DR N-PSDB; AAD07291.
XX
XX Novel murine serine/threonine kinase polypeptides and polynucleotides
PT for identifying agonists, antagonists useful for treating anaemia,
PT cancer, rheumatoid arthritis, psoriasis, psychotic and neurological
PT disorders
XX
XX Claim 2; Page 24; 34pp; English.
XX
XX The present sequence is mouse serine/threonine kinase, YAK3.
CC The YAK3 is useful for screening compounds which inhibit or stimulate
CC the function of YAK3. The identified agonists and antagonists are useful
CC in treatment of human diseases including neutropenia, cytopenia,
CC anaemias including anaemias due to renal insufficiency or due to chronic
CC disease such as autoimmunity or cancer, polycythaemia, myelosuppression,
CC infertility, testicular cancer, bone loss including osteoporosis,
CC inflammatory diseases such as adult respiratory disease syndrome,
CC rheumatoid arthritis, psoriasis, dermatitis, asthma, allergies,
CC infections such as bacterial, fungal, protozoan and viral infections,
CC particularly infections caused by HIV and other immunodeficiency
CC disorders, pain, injury, anorexia, Parkinson's disease, cardiovascular

CC diseases including atherosclerosis, hypotension, hypertension, myocardial
 CC infarction, ulcers, psychotic and neurological disorders including
 CC anxiety, schizophrenia, dementia, delirium and dyskinesias such as
 CC Huntington's disease. The YAK3 is also used as a probe to screen suitable
 CC libraries to obtain murine YAK3 genomic DNA clone and for chromosome
 CC localisation studies.

SQ Sequence 551 AA;

Query Match 50.5%; Score 53; DB 22; Length 551;

Best Local Similarity 64.3%; Pred. No. 9.2;

Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 KQYEQHLTDYEKIK 17

|||: ||| |||::

Db 112 kqykhltayekle 125

RESULT 11

AA065662
 ID AAB65662 standard; Protein; 553 AA.

XX AC AAB65662;

XX 27-MAR-2001 (first entry)

DE Novel protein kinase, SEQ ID NO: 189.

XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
 KW dermatological; antidiabetic; antinfertility; gene therapy; vaccine;
 KW immune disorder; cardiovascular disease; neurodegenerative disease;
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis.

XX Homo sapiens.

OS WO200073469-A2.

PN 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US14842.

XX 28-MAY-1999; 99US-0136503.

XX (SUGE-) SUGEN INC.

XX Plowman GD, Martinez R, Whyte D, Sudersanam S;

DR WPI: 2001-032161/04.

DR N-PSDB; AAF44689.

XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers -

PS Claim 10; Fig 1; 310pp; English.

XX The present sequence is a novel protein kinase. The novel protein kinases
 CC and the nucleic acids that encode them may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC and kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomyopathies, strokes, renal failure,
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,

CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 CC reproductive disorders.

SQ Sequence 553 AA;

Query Match 50.5%; Score 53; DB 22; Length 553;

Best Local Similarity 64.3%; Pred. No. 9.2;

Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 KQYEQHLTDYEKIK 17

|||: ||| |||::

Db 113 kqykhltayekle 126

RESULT 12

AA075791
 ID AAW75791 standard; Protein; 568 AA.

XX AC AAW75791;

XX 21-DEC-1998 (first entry)

XX Human protein kinase hYAK3-beta.

XX Protein kinase; hYAK-beta; human; bone loss; osteoporosis;
 KW inflammation; acute respiratory distress syndrome; osteoarthritis;
 KW rheumatoid arthritis; inflammatory bowel disease; psoriasis;
 KW dermatitis; asthma; allergy; infection; HIV; immunodeficiency;
 KW cachexia; septic shock; pain; injury; cancer; testicular cancer;
 KW anorexia, bulimia; Parkinson's disease; cardiovascular disease;
 KW restenosis; atherosclerosis; heart failure; myocardial infarction;
 KW hypotension; hypertension; urinary retention; angina pectoris;
 KW ulcer; benign prostatic hypertrophy; psychosis; schizophrenia;
 KW anxiety; manic depression; delirium; severe mental retardation;
 KW dementia; Huntington's disease; Gilles de la Tourette syndrome;
 KW dyskinesia; diagnosis; therapy; vaccine.

XX Homo sapiens.

XX EP870825-A1.

XX 14-OCT-1998.

XX 05-MAR-1998; 98EP-0301641.

XX 07-APR-1997; 97US-0835170.

XX 05-MAR-1997; 97US-0040618.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX Creasy CL, Xie W;

XX WPI: 1998-523155/45.

XX N-PSDB; AAV57440.

XX New DNA encoding hYAK3 human protein kinase polypeptides - used to
 PT treat and diagnose e.g. osteoporosis, inflammatory diseases,
 PT inflammatory bowel disease, psoriasis, dermatitis, asthma,
 PT allergies, infections, septic shock, anorexia and cancer

PS Claim 15; Fig 2; 23pp; English.

XX This is the amino acid sequence of hYAK-beta, a novel human
 CC protein kinase that is expressed predominantly in testis tissue.
 CC The sequence was deduced from a full-length cDNA clone (see AAV57440)
 CC isolated from a testis cDNA library. Another clone (see AAV57439),
 CC from a skeletal muscle cDNA library, codes for a longer form (see
 CC AAV5790), termed hYAK-alpha. This may be a result of alternative
 CC splicing. The invention provides host cells and methods for
 CC producing hYAK3 polypeptides, as well as methods for treating
 CC subjects having need to enhance or inhibit hYAK3 activity, for

diagnosing a disease related to expression or activity of hYAK3,
and methods for identifying agonist and antagonist compounds. hYAK3
polypeptides and polynucleotides can be used to treat and diagnose
bone loss (e.g. osteoporosis), inflammatory diseases (e.g. ARDS),
rheumatoid arthritis, osteoarthritis, inflammatory bowel disease,
psoriasis, dermatitis, asthma, allergies, infections (e.g. HIV),
bacterial, fungal, protozoal and viral infections such as HIV),
HIV-associated cachexia and other immunodeficiency disorders,
septic shock, pain, injury, cancers (e.g. testicular cancer),
anorexia, bulimia, Parkinson's disease, cardiovascular disease
(e.g. restenosis, atherosclerosis, acute heart failure and
myocardial infarction), hypotension, hypertension, urinary
retention, angina pectoris, ulcers, benign prostatic hypertrophy
and psychotic and neurological disorders (e.g. anxiety,
schizophrenia, manic depression, delirium, dementia, severe mental
retardation and dyskinesias such as Huntington's disease or Gilles
de la Tourette syndrome).

XX SQ Sequence 568 AA;

Query Match 50.5%; Score 53; DB 19; Length 568;
Best Local Similarity 64.3%; Pred. No. 9.5;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 4 KOYEOHLTDYEKIK 17
| | | | | | | | | |
Db 128 kqkhhltayekle 141

RESULT 13
AAB00059
ID AAB00059 standard; Protein; 568 AA.
AC AAB00059;
XX
DT 08-NOV-2000 (first entry)
DE hYAK3-2 polypeptide.
KW Human YAK3-2; serine/threonine kinase; kinase;
KW phosphorylation; signal transduction; antisense; suppression;
KW bone marrow; foetal liver; testis; prevention; treatment;
KW diagnosis; anaemia; neutropenia; cytopaenia; polycythaemia;
KW myelosuppression; bone loss; osteoporosis; inflammatory diseases;
KW adult respiratory disease syndrome; ARDS; rheumatoid arthritis;
KW osteoarthritis; inflammatory bowel disease; IBD; psoriasis;
KW dermatitis; asthma; allergies; infection;
KW human immunodeficiency virus; HIV; HIV-associated cachexia;
KW septic shock; pain; injury; cancer; anorexia; bulimia;
KW Parkinson's disease; cardiovascular disease; restenosis;
KW atherosclerosis; acute heart failure; myocardial infarction;
KW hypotension; hypertension; urinary retention; anginal pectoris;
KW ulcers; benign prostatic hypertrophy.

OS Homo sapiens.
XX
XX WO200044941-A1.
XX
XX PD 03-AUG-2000.
XX
XX PF 01-FEB-2000; 2000WO-US02544.
XX
XX PR 01-FEB-1999; 99US-0118045.
XX PR 10-AUG-1999; 99US-0371674.
XX PR 15-NOV-1999; 99US-0440236.
XX
XX PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
XX PI Lord KA, Dillon SB, Creasy C;
XX
XX DR WPI: 2000-499340/44.
XX DR N-PSDB: AAA47745.

XX hYAK3-2 protein serine kinases polypeptides and nucleic acids useful
PT for treating e.g. anemia, neutropenia and cytopenia
XX
PS Claim 12; Page 40-41; 45pp; English.
XX
CC A number of polypeptide growth factors and hormones mediate their
CC cellular effects through a signal transduction pathway. Transduction
CC of signals from the cell surface receptors to intracellular effectors
CC frequently involves phosphorylation or dephosphorylation of specific
CC protein substrates by regulatory protein serine/threonine kinases
CC (PSTK) and phosphatases. Aberrant protein serine/threonine kinase
CC activity has been implicated in a number of pathologies and disease
CC states, thus, these enzymes and their pathways provide suitable
CC targets for drug design. YAK1 is a PSTK with sequence homology to
CC CDK's (cyclin dependent kinases, which have a role in cell cycling).
CC It was originally identified in yeast where it can act as a cell
CC cycle suppressor. Novel human homologues of yeast YAK1 designated
CC YAK3-2 have been described. YAK3-2 is expressed predominantly in
CC haematopoietic tissues such as bone marrow, foetal liver and in the
CC testis. Inhibitors of hYAK3-2 are expected to stimulate expression of
CC cells in which it is expressed. The use of antisense oligonucleotides
CC to achieve such ends is described. Nucleic acids encoding the hYAK3-2
CC and the hYAK3-2 protein may be used in the prevention, treatment and
CC diagnosis of diseases associated with inappropriate hYAK3-2
CC expression. Disorders that may be treated using the nucleic acids and
CC polypeptides include haematopoietic cellular deficiency (either
CC anaemia, neutropenia, cytopaenia, polycythaemia, myelosuppression,
CC inflammatory diseases (e.g. adult respiratory disease syndrome,
CC ARDS)), rheumatoid arthritis, osteoarthritis, inflammatory bowel
CC disease (IBD), psoriasis, dermatitis, asthma, allergies, infections
CC (e.g. bacterial, fungal, protozoan and viral infections, particularly
CC infections caused by human immunodeficiency virus (HIV)-1 or HIV-2,
CC HIV-associated cachexia and other immunodeficiency disorders),
CC septic shock, pain, injury, cancers (e.g. testicular cancer),
CC anorexia, bulimia, Parkinson's disease, cardiovascular disease
CC (e.g. restenosis, atherosclerosis, acute heart failure and myocardial
CC infarction), hypotension, hypertension, urinary retention, anginal
CC pectoris, ulcers, benign prostatic hypertrophy and psychotic and
CC neurological disorders.
XX SQ Sequence 568 AA;

Query Match 50.5%; Score 53; DB 21; Length 568;
Best Local Similarity 64.3%; Pred. No. 9.5;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 4 KOYEOHLTDYEKIK 17
| | | | | | | | | |
Db 128 kqkhhltayekle 141

RESULT 14
AAE02013
ID AAE02013 standard; Protein; 568 AA.
XX
XX AC AAE02013;
XX
XX DT 31-JUL-2001 (first entry)
XX
XX DE Human YAK3b (hYAK3b) serine/threonine protein kinase (STPK).
XX
XX KW Human; YAK3b; serine/threonine protein kinase; STPK; signal transduction;
KW therapy; bone loss; osteoporosis; inflammatory disease; osteoarthritis;
KW adult respiratory disease syndrome; ARDS; rheumatoid arthritis; allergy;
KW inflammatory bowel disease; IBD; psoriasis; dermatitis; asthma; cachexia;
KW human immuno deficiency virus; HIV; immunodeficiency disorder; psychotic;
KW septic shock; restenosis; pain; cancer; anorexia; bulimia; hypotension;
KW atherosclerosis; Parkinson's disease; cardiovascular disease; anxiety;
KW acute heart failure; myocardial infarction; hypertension; schizophrenia;
KW urinary retention; angina pectoris; ulcer; benign prostatic hypertrophy;

KW neurological disorder; depression; delirium; dementia;
KW Huntington's disease; mental retardation; proliferative disease;
KW Gilles de la Tourette's syndrome;
XX Homo sapiens.
XX WO200132918-A1.
XX 10-MAY-2001.
XX 06-NOV-2000; 2000WO-US30542.
XX 05-NOV-1999; 99US-0163901.
XX 03-NOV-2000; 2000US-0706426.
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX (SMIK) SMITHKLINE BEECHAM PLC.
XX Kassiss S, Annan RS, Lee JC;
XX WPI; 2001-335837/35.
XX N-PSDB; AAD06043.
XX Screening for YAK polypeptide kinase activity comprising contact with a
XX candidate peptide to detect molecules useful for treating disorders
XX including the treatment of bone loss -
XX
XX Claim 2; Page 29-30; 49pp; English.
XX The present sequence is human YAK3b (HYAK3b) serine/threonine protein
XX kinase (STPK). YAK polypeptides are involved in the signal transduction
XX pathways. YAK modulators are used to treat bone loss such as
XX osteoporosis, inflammatory diseases such as adult respiratory disease
XX syndrome (ARDS), rheumatoid arthritis, osteoarthritis, inflammatory bowel
XX disease (IBD), psoriasis, dermatitis, asthma and allergies, bacterial,
XX fungal and viral infections such as human immunodeficiency virus-1
XX (HIV-1) and HIV-2 infections, HIV-associated cachexia and other
XX immunodeficiency disorders, septic shock, pain, injury, cancers,
XX anorexia, bulimia, Parkinson's disease, cardiovascular diseases such as
XX restenosis, atherosclerosis, acute heart failure, myocardial infarction,
XX hypotension, hypertension, urinary retention, angina pectoris, ulcers,
XX benign prostatic hypertrophy and psychotic and neurological disorders
XX such as anxiety, schizophrenia, manic depression, delirium, dementia,
XX severe mental retardation and dyskinesias such as Huntington's disease or
XX Gilles de la Tourette's syndrome and proliferative diseases.
XX
XX Sequence 568 AA;

XX Query Match 50.5%; Score 53; DB 22; Length 568;
XX Best Local Similarity 64.3%; Pred. No. 9.5;
XX Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 KOYEQHLTDYEKIK 17
DB 128 kqykhltayekie 141

RESULT 15
AAW75790
ID AAW75790 standard; Protein; 588 AA.
XX
XX AAW75790;
XX
XX 21-DEC-1998 (first entry)
XX
XX Human protein kinase HYAK3-alpha.
XX
KW Protein kinase; HYAK-alpha; human; bone loss; osteoporosis;
KW inflammation; acute respiratory distress syndrome; osteoarthritis;
KW rheumatoid arthritis; inflammatory bowel disease; psoriasis;
KW dermatitis; asthma; allergy; infection; HIV; immunodeficiency;
KW cachexia; septic shock; pain; injury; cancer; testicular cancer;

KW anorexia, bulimia; Parkinson's disease; cardiovascular disease;
KW restenosis; atherosclerosis; heart failure; myocardial infarction;
KW hypotension; hypertension; urinary retention; angina pectoris;
KW ulcer; benign prostatic hypertrophy; psychosis; schizophrenia;
KW anxiety; manic depression; delirium; severe mental retardation;
KW dementia; Huntington's disease; Gilles de la Tourette syndrome;
KW dyskinesia; diagnosis; therapy; vaccine.
XX Homo sapiens.
XX
XX EP870825-A1.
XX 14-OCT-1998.
XX 05-MAR-1998; 98EP-0301641.
XX 07-APR-1997; 97US-0835170.
XX 05-MAR-1997; 97US-0040618.
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX Creasy CL, Xie W;
XX WPI; 1998-523155/45.
XX N-PSDB; AAV57439.
XX New DNA encoding HYAK3 human protein kinase polypeptides - used to
XX treat and diagnose e.g. osteoporosis, inflammatory diseases,
XX inflammatory bowel disease, psoriasis, dermatitis, asthma,
XX allergies, infections, septic shock, anorexia and cancer
XX Claim 15; Fig 1; 23pp; English.
XX This is the amino acid sequence of HYAK-alpha, a novel human
XX protein kinase that is expressed predominantly in skeletal muscle.
XX The sequence was deduced from a full-length cDNA clone (see AAV57439)
XX isolated from a skeletal muscle cDNA library. Another clone (see
XX AAV57440), from a testis cDNA library, codes for a shorter form (see
XX AAW5791), termed HYAK-beta. This may be a result of alternative
XX splicing. The invention provides host cells and methods for
XX producing HYAK3 polypeptides, as well as methods for treating
XX subjects having need to enhance or inhibit HYAK3 activity, for
XX diagnosing a disease related to expression or activity of HYAK3,
XX and methods for identifying agonist and antagonist compounds. HYAK3
XX polypeptides and polynucleotides can be used to treat and diagnose
XX bone loss (e.g. osteoporosis), inflammatory diseases (e.g. ARDS),
XX rheumatoid arthritis, osteoarthritis, inflammatory bowel disease,
XX psoriasis, dermatitis, asthma, allergies, infections (e.g.
XX bacterial, fungal, protozoal and viral infections such as HIV),
XX HIV-associated cachexia and other immunodeficiency disorders,
XX septic shock, pain, injury, cancers (e.g. testicular cancer),
XX anorexia, bulimia, Parkinson's disease, cardiovascular disease
XX (e.g. restenosis, atherosclerosis, acute heart failure and
XX myocardial infarction), hypotension, hypertension, urinary
XX retention, angina pectoris, ulcers, benign prostatic hypertrophy
XX and psychotic and neurological disorders (e.g. anxiety,
XX schizophrenia, manic depression, delirium, dementia, severe mental
XX retardation and dyskinesias such as Huntington's disease or Gilles
XX de la Tourette syndrome).
XX
XX Sequence 588 AA;

Query Match 50.5%; Score 53; DB 19; Length 588;
Best Local Similarity 64.3%; Pred. No. 9.8;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 KOYEQHLTDYEKIK 17
DB 148 kqykhltayekie 161

Search completed: January 29, 2002, 10:21:50
Job time: 426 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:24:09 : Search time 133.18 Seconds
(without alignments)
3.210 Million cell updates/sec

Title: US-09-763-397A-18

Perfect score: 105

Sequence: 1 DQPKQYEQHLTDYKKEG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PC1US_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	100.0	628	1	US-08-257-073-9
2	53	50.5	568	2	US-08-835-170-4
3	53	50.5	568	4	US-09-359-257-4
4	53	50.5	588	2	US-08-835-170-2
5	53	50.5	588	4	US-09-359-257-2
6	44	41.9	308	2	US-08-897-340-33
7	44	41.9	308	4	US-09-252-329-33
8	44	41.9	458	1	US-08-336-618-24
9	44	41.9	459	1	US-08-336-618-12
10	44	41.9	459	1	US-08-336-618-26
11	44	41.9	484	2	US-08-879-260-4
12	44	41.9	484	3	US-09-231-529-4
13	44	41.9	484	4	US-08-977-816-4
14	44	41.9	1865	1	US-08-588-985-2
15	44	41.9	1865	1	US-08-971-988-2
16	43	41.0	1349	1	US-08-612-734B-2
17	43	41.0	1588	5	PCT-US93-07261-11
18	43	41.0	1663	5	PCT-US93-07261-16
19	41	39.0	251	4	US-09-268-364-16
20	41	39.0	770	2	US-08-209-521-13
21	41	39.0	770	4	US-08-961-810-123
22	41	39.0	770	4	US-08-352-902D-123
23	41	39.0	1334	2	US-08-996-545-2
24	41	39.0	1334	4	US-09-328-320-2
25	40	38.1	438	2	US-08-897-340-34
26	40	38.1	438	4	US-09-252-329-34
27	39	37.1	236	4	US-08-961-083-96

28 39 37.1 338 3 US-08-722-184-4 Sequence 4, Appli
29 39 37.1 338 4 US-09-043-937A-8 Sequence 8, Appli
30 39 37.1 1872 1 US-08-188-582-14 Sequence 14, Appl
31 39 37.1 1872 1 US-08-646-715-14 Sequence 14, Appl
32 39 37.1 1893 1 US-08-188-582-11 Sequence 11, Appl
33 39 37.1 1893 1 US-08-646-715-11 Sequence 11, Appl
34 38.5 36.7 617 4 US-09-314-242-2 Sequence 2, Appli
35 38.5 36.7 1026 1 US-07-998-003A-95 Sequence 95, Appl
36 38.5 36.7 1026 1 US-08-453-274B-95 Sequence 95, Appl
37 38.5 36.7 1026 1 US-08-453-695A-95 Sequence 95, Appl
38 38.5 36.7 1026 1 US-08-268-161A-95 Sequence 95, Appl
39 38.5 36.7 1026 2 US-08-453-702A-95 Sequence 95, Appl
40 38.5 36.7 1026 4 US-09-099-639-95 Sequence 95, Appl
41 38.5 36.7 1026 5 PCT-US93-12588-95 Sequence 95, Appl
42 38.5 36.7 1026 5 PCT-US95-08071-95 Sequence 95, Appl
43 38.5 36.7 1203 1 US-07-998-003A-103 Sequence 103, App
44 38.5 36.7 1203 1 US-08-453-274B-103 Sequence 103, App
45 38.5 36.7 1203 1 US-08-453-695A-103 Sequence 103, App

ALIGNMENTS

RESULT 1
US-08-257-073-9
; Sequence 9, Application US/08257073
; Patent No. 5766597
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: de Taisne, Charles
; APPLICANT: Tine, John A.
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtiss, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue, 25th Floor
; CITY: New York
; STATE: New York
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,073
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/075,783
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/852,305
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,183
; FILING DATE: 20-MAR-1991

ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2570

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 628 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

```
; FRAGMENT TYPE: internal
US-08-257-073-9

Query Match      100.0%; Score 105; DB 1; Length 628;
Best Local Similarity 100.0%; Pred. No. 9,9e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DDPKQYEQHLTDYEKIK 19
   |||||
Db 348 DDPKQYEQHLTDYEKIK 366

RESULT 2
US-08-835-170-4
; Sequence 4, Application US/08835170
; Patent No. 5965420
; GENERAL INFORMATION:
; APPLICANT: Creasy, Caretha
; TITLE OF INVENTION: Human Protein Kinases HYAK3
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,170
; FILING DATE: 26-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: GH50004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 568 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-835-170-4

Query Match      50.5%; Score 53; DB 2; Length 568;
Best Local Similarity 64.3%; Pred. No. 1.2;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 KQYEQHLTDYEKIK 17
   |||
Db 128 KQYKHLTAYEKL 141

RESULT 3
US-09-359-257-4
; Sequence 4, Application US/09359257
; Patent No. 6165766
; GENERAL INFORMATION:
; APPLICANT: Creasy, Caretha
; TITLE OF INVENTION: Human Protein Kinases HYAK3
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,170
; FILING DATE: 26-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: GH50004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 568 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-835-170-4

Query Match      50.5%; Score 53; DB 2; Length 568;
Best Local Similarity 64.3%; Pred. No. 1.2;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 KQYEQHLTDYEKIK 17
   |||
Db 128 KQYKHLTAYEKL 141

RESULT 4
US-08-835-170-2
; Sequence 2, Application US/08835170
; Patent No. 5965420
; GENERAL INFORMATION:
; APPLICANT: Creasy, Caretha
; TITLE OF INVENTION: Human Protein Kinases HYAK3
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,170
; FILING DATE: 26-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: GH50004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 588 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-835-170-2
```

Query Match 50.5%; Score 53; DB 2; Length 588;
Best Local Similarity 64.3%; Pred. No. 1.2;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 4 KOYEQHLDYKIK 17
Db 148 KOYKHLTAYEKL 161

RESULT 5

US-09-359-257-2
; Sequence 2, Application US/09359257
; Patent No. 6165766
; GENERAL INFORMATION:
; APPLICANT: Creasy, Caretha
; TITLE OF INVENTION: HUMAN PROTEIN KINASES HYAK3
; FILE REFERENCE: GH50004X1D1
; CURRENT APPLICATION NUMBER: US/09/359,257
; CURRENT FILING DATE: 1999-07-22
; EARLIER APPLICATION NUMBER: 08/835,170
; EARLIER FILING DATE: 1997-04-07
; EARLIER APPLICATION NUMBER: 60/040618
; EARLIER FILING DATE: 1997-03-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Human
US-09-359-257-2

Query Match 50.5%; Score 53; DB 4; Length 588;
Best Local Similarity 64.3%; Pred. No. 1.2;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 4 KOYEQHLDYKIK 17
Db 148 KOYKHLTAYEKL 161

RESULT 6

US-08-897-340-33
; Sequence 33, Application US/08897340
; Patent No. 5955306
; GENERAL INFORMATION:
; APPLICANT: Gimeno, Carlos J. and Errada, Patrick, R.
; TITLE OF INVENTION: Weight Control Pathway Genes and Uses
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/897,340
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/715,032
; FILING DATE: 17-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Silveri, Jean M.

; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: MNI-005CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 308 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-897-340-33

Query Match 41.9%; Score 44; DB 2; Length 308;
Best Local Similarity 53.8%; Pred. No. 16;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 4 KOYEQHLDYKIK 16
Db 270 EOYEEAVRDYK 282

RESULT 7

US-09-252-329-33
; Sequence 33, Application US/09252329
; Patent No. 6147192
; GENERAL INFORMATION:
; APPLICANT: Gimeno, Carlos J. and Errada, Patrick, R.
; TITLE OF INVENTION: Weight Control Pathway Genes and Uses
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/252,329
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/897,340
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silveri, Jean M.
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: MNI-005CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 308 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-09-252-329-33

Query Match 41.9%; Score 44; DB 4; Length 308;
Best Local Similarity 53.8%; Pred. No. 16;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 KYEQHLTDYEKI 16
:||||:||||:
Db 270 EQVEEAVRDYEK 282

RESULT 8

US-08-336-618-24

; Sequence 24, Application US/08336618

; Patent No. 5763590

; GENERAL INFORMATION:

; APPLICANT: Peattie, Debra A.

; APPLICANT: Harding, Matthew W.

; APPLICANT: Livingston, David J.

; TITLE OF INVENTION: ISOLATION OF AN MR 52,000 FK506 BINDING

; TITLE OF INVENTION: PROTEIN AND MOLECULAR CLONING OF A CORRESPONDING HUMAN

; NUMBER OF SEQUENCES: 32

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hamilton, Brook, Smith and Reynolds, P.C.

; STREET: Two Militia Drive

; CITY: Lexington

; STATE: Massachusetts

; COUNTRY: U.S.A.

; ZIP: 02173

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/336,618

; FILING DATE: 09-NOV-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/963,325

; FILING DATE: 16-OCT-1992

; APPLICATION NUMBER: US 07/777,752

; FILING DATE: 11-OCT-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/

; FILING DATE: 09-OCT-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Granahan, Patricia

; REGISTRATION NUMBER: 32,227

; REFERENCE/DOCKET NUMBER: VPI91-06A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-861-6240

; TELEFAX: 617-861-9540

; INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 458 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-336-618-24

Query Match

Best Local Similarity 41.9%; Score 44; DB 1; Length 458;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 QYEQHLTDYEKI 18

:|||:|||:||||

Db 244 KYEVLHLSPEKAKE 257

RESULT 9

US-08-336-618-12

; Sequence 12, Application US/08336618

; Patent No. 5763590

; GENERAL INFORMATION:

; APPLICANT: Peattie, Debra A.

; APPLICANT: Harding, Matthew W.

; APPLICANT: Livingston, David J.
; TITLE OF INVENTION: ISOLATION OF AN MR 52,000 FK506 BINDING
; TITLE OF INVENTION: PROTEIN AND MOLECULAR CLONING OF A CORRESPONDING HUMAN
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hamilton, Brook, Smith and Reynolds, P.C.

; STREET: Two Militia Drive

; CITY: Lexington

; STATE: Massachusetts

; COUNTRY: U.S.A.

; ZIP: 02173

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/336,618

; FILING DATE: 09-NOV-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/963,325

; FILING DATE: 16-OCT-1992

; APPLICATION NUMBER: US 07/777,752

; FILING DATE: 11-OCT-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/

; FILING DATE: 09-OCT-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Granahan, Patricia

; REGISTRATION NUMBER: 32,227

; REFERENCE/DOCKET NUMBER: VPI91-06A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-861-6240

; TELEFAX: 617-861-9540

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 459 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-336-618-12

Query Match

Best Local Similarity 41.9%; Score 44; DB 1; Length 459;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 QYEQHLTDYEKI 18

:|||:|||:||||

Db 244 KYEVLHLSPEKAKE 257

RESULT 10

US-08-336-618-26

; Sequence 26, Application US/08336618

; Patent No. 5763590

; GENERAL INFORMATION:

; APPLICANT: Peattie, Debra A.

; APPLICANT: Harding, Matthew W.

; APPLICANT: Livingston, David J.

; TITLE OF INVENTION: ISOLATION OF AN MR 52,000 FK506 BINDING

; TITLE OF INVENTION: PROTEIN AND MOLECULAR CLONING OF A CORRESPONDING HUMAN

; NUMBER OF SEQUENCES: 32

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hamilton, Brook, Smith and Reynolds, P.C.

; STREET: Two Militia Drive

; CITY: Lexington

; STATE: Massachusetts

; COUNTRY: U.S.A.

; ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,618
FILING DATE: 09-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/963,325
FILING DATE: 16-OCT-1992
APPLICATION NUMBER: US 07/777,752
FILING DATE: 11-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/
FILING DATE: 09-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: VPI91-06A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-336-618-26

Query Match 41.9%; Score 44; DB 1; Length 459;
Best Local Similarity 57.1%; Pred. No. 24;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 QYEQHLTDYEKI 18
DB 244 KYELHLKSEKAKE 257

RESULT 11
US-08-879-260-4
Sequence 4, Application US/08879260
Patent No. 5935851
GENERAL INFORMATION:
APPLICANT: Murthy, Anita E.
APPLICANT: Gusella, James F.
TITLE OF INVENTION: TPR-Containing Genes
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C
STREET: 1100 New York Ave, N.W., Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/879,260
FILING DATE: 19JUN1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/020,204
FILING DATE: 20JUN1996
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609.4260001/JAG/SRL
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 484 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-879-260-4

Query Match 41.9%; Score 44; DB 2; Length 484;
Best Local Similarity 53.8%; Pred. No. 26;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 QYEQHLTDYEKI 16
DB 332 EQYEAARDYEV 344

RESULT 12
US-09-231-529-4
Sequence 4, Application US/09231529
Patent No. 6096308
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN PROTEIN KINASE AND KINASE INHIBITORS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/231,529
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/977,816
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0429 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-853-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 484 amino acids
TYPE: amino acid
STRADEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLNNOT11
CLONE: 701698
US-09-231-529-4

Query Match 41.9%; Score 44; DB 3; Length 484;

Best Local Similarity 53.8%; Pred. No. 26;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 KYEQHLTDYEKI 16
:||||:||||:
Db 332 EQYEAVRDYEKV 344

RESULT 13

US-08-977-816-4
; Sequence 4, Application US/08977816
; Patent No. 6194186
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN PROTEIN KINASE AND KINASE INHIBITORS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/977,816
; FILING DATE: Filed Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0429 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 484 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: COLNOT11
; CLONE: 701698
; US-08-977-816-4

Query Match 41.9%; Score 44; DB 4; Length 484;
Best Local Similarity 53.8%; Pred. No. 26;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 KYEQHLTDYEKI 16
:||||:||||:
Db 332 EQYEAVRDYEKV 344

RESULT 14

US-08-588-985-2
; Sequence 2, Application US/08588985
; Patent No. 5777094
; GENERAL INFORMATION:
; APPLICANT: Michiyuki MATSUDA et al.

; TITLE OF INVENTION: cDNA OF DOCK180 GENE AND DOCK180 PROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/588,985
; FILING DATE: January 19, 1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1865 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE: spleen cell of homo sapiens
; US-08-588-985-2

Query Match 41.9%; Score 44; DB 1; Length 1865;
Best Local Similarity 46.7%; Pred. No. 11e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 KYEQHLTDYEKI 18
:||||:||||:
Db 1302 EQYENMFQYQLSE 1316

RESULT 15

US-08-971-988-2
; Sequence 2, Application US/08971988
; Patent No. 5786461
; GENERAL INFORMATION:
; APPLICANT: Michiyuki MATSUDA et al.
; TITLE OF INVENTION: cDNA OF DOCK180 GENE AND DOCK180 PROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/971,988
; FILING DATE: 17-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/588,985
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1865 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE: spleen cell of homo sapiens
US-08-971-988-2

```

```

Query Match      41.9%  Score 44; DB 1; Length 1865;
Best Local Similarity 46.7%; Pred. NO. 1.le+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy      4 KQEQHLDYKIKE 18
Db      1302 EQYENEMFDYQLSE 1316

```

Search completed: January 29, 2002, 10:24:10
Job time: 516 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:26:44 ; Search time 144.96 Seconds
(without alignments)
9.984 Million cell updates/sec

Title: US-09-763-397A-18

Perfect score: 105

Sequence: 1 DQPKQYEQHLDYKIKEG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	100.0	622	2 D44986	apical membrane an
2	105	100.0	622	2 B44986	apical membrane an
3	105	100.0	622	2 A32499	apical membrane an
4	105	100.0	622	2 A44986	apical membrane an
5	105	100.0	622	2 C44986	apical membrane an
6	81	77.1	563	2 A44944	apical membrane an
7	79	75.2	563	2 A39238	66K merozoite surf
8	72	68.6	558	2 A44984	apical membrane an
9	52	49.5	463	2 T31570	hypothetical prote
10	51	48.6	333	2 C81263	hypothetical prote
11	47	44.8	372	2 C81263	probable integral
12	47	44.8	1165	2 I40644	botulinum neurotox
13	46	43.8	222	2 E84486	hypothetical prote
14	46	43.8	319	2 T35538	cytochrome-c oxid
15	46	43.8	696	2 A70431	ribonucleotide red
16	46	43.8	1573	2 S01845	DNA (cytosine-5-)
17	45	42.9	243	2 I51746	MHC class II alpha
18	45	42.9	332	2 H69497	ketol-acid reducto
19	45	42.9	363	2 JE0111	lectin-like oxidiz
20	45	42.9	621	2 E82912	hypothetical prote
21	45	42.9	1022	2 T17406	developmental prot
22	44	41.9	256	2 H64355	hypothetical prote
23	44	41.9	281	2 F75216	hypothetical prote
24	44	41.9	299	2 A83700	hypothetical prote
25	44	41.9	458	1 A42386	hsp 90-binding pro
26	44	41.9	459	2 A46372	immunophilin FRBP5
27	44	41.9	986	2 T10754	cis-Golgi matrix p
28	44	41.9	1300	2 T18364	ro-3 protein - Neu
29	43.5	41.4	70	2 E82895	hypothetical prote

30	43.5	41.4	320	2 T27372	hypothetical prote
31	43.5	41.4	329	2 T19872	hypothetical prote
32	43.5	41.4	441	2 T43544	sat1 protein - fis
33	43.5	41.4	550	2 T39942	sat1p - fission ye
34	43	41.0	169	2 H70377	hypothetical prote
35	43	41.0	279	2 D71453	hypothetical prote
36	43	41.0	289	2 B72099	conserved hypothet
37	43	41.0	289	2 H86522	CT144 hypothetical
38	43	41.0	307	2 F86865	conserved hypothet
39	43	41.0	332	2 H75044	ketol-acid reducto
40	43	41.0	337	2 D84079	hypothetical prote
41	43	41.0	363	2 H69393	iron-sulfur cluste
42	43	41.0	394	2 B82931	hypothetical membr
43	43	41.0	516	1 PIWL5	L1 protein - human
44	43	41.0	614	2 D82942	glucose inhibited
45	43	41.0	765	2 E96558	hypothetical prote

ALIGNMENTS

RESULT 1

D44986

apical membrane antigen 1 - malaria parasite (Plasmodium falciparum) (strain 768)
C:Species: Plasmodium falciparum

C>Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000

C:Accession: D44986

R:Thomas, A.W.; Waters, A.P.; Carr, D.

Mol. Biochem. Parasitol. 42, 285-288, 1990

A:Title: Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate a
A:Reference number: A44986; MUID:91101665

A:Accession: D44986

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: DNA

A:Residues: 1-622 <THO>

A:Cross-references: GB:M34555

C:Keywords: membrane protein; surface antigen

Query Match 100.0%; Score 105; DB 2; Length 622;

Best Local Similarity 100.0%; Pred. No. 3.7e-08;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DQPKQYEQHLDYKIKEG 19

|||||

Db 348 DQPKQYEQHLDYKIKEG 366

RESULT 2

B44986

apical membrane antigen 1 - malaria parasite (Plasmodium falciparum) (strain Thai Tn)
C:Species: Plasmodium falciparum

C>Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000

C:Accession: B44986

R:Thomas, A.W.; Waters, A.P.; Carr, D.

Mol. Biochem. Parasitol. 42, 285-288, 1990
A:Title: Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate a
A:Reference number: A44986; MUID:91101665

A:Accession: B44986

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: DNA

A:Residues: 1-622 <THO>

A:Cross-references: GB:M34553

C:Keywords: membrane protein; surface antigen

Query Match 100.0%; Score 105; DB 2; Length 622;

Best Local Similarity 100.0%; Pred. No. 3.7e-08;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DQPKQYEQHLDYKIKEG 19

|||||

Db 348 DQPKQYEQHLDYKIKEG 366

```
RESULT 3
A32499 apical membrane antigen 1 - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 12-Oct-1989 #sequence_revision 12-Oct-1989 #text_change 09-Jun-2000
C:Accession: A32499
R:Peterson, M.G.; Marshall, V.M.; Smythe, J.A.; Crewther, P.E.; Lew, A.; Silva, A.; Ande
Mol. Cell. Biol. 9, 3151-3154, 1989
A:Title: Integral membrane protein located in the apical complex of Plasmodium falcipar
A:Reference number: A32499; MUID:89384584
A:Accession: A32499
A:Status: preliminary
A:Molecule type: DNA; mRNA
A:Residues: 1-622 <PET>
A:Cross-references: GB:M27133; NID:gi60072; PID:gi60073
C:Keywords: membrane protein; surface antigen

Query Match 100.0%; Score 105; DB 2; Length 622;
Best Local Similarity 100.0%; Pred. No. 3.7e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DDPKQYEQHLTDYEKIQEG 19
Db 348 DDPKQYEQHLTDYEKIQEG 366

RESULT 4
A44986 apical membrane antigen 1 - malaria parasite (Plasmodium falciparum) (strain CAMP)
C:Species: Plasmodium falciparum
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000
C:Accession: A44986
R:Thomas, A.W.; Waters, A.P.; Carr, D.
Mol. Biochem. Parasitol. 42, 285-288, 1990
A:Title: Analysis of variation in PR83, an erythrocytic merozoite vaccine candidate anti
A:Reference number: A44986; MUID:91101665
A:Accession: A44986
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: DNA
A:Residues: 1-622 <THO>
A:Cross-references: GB:M34552
C:Keywords: membrane protein; surface antigen

Query Match 100.0%; Score 105; DB 2; Length 622;
Best Local Similarity 100.0%; Pred. No. 3.7e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DDPKQYEQHLTDYEKIQEG 19
Db 348 DDPKQYEQHLTDYEKIQEG 366

RESULT 5
C44986 apical membrane antigen 1 - malaria parasite (Plasmodium falciparum) (strain FCR 3)
C:Species: Plasmodium falciparum
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000
C:Accession: C44986
R:Thomas, A.W.; Waters, A.P.; Carr, D.
Mol. Biochem. Parasitol. 42, 285-288, 1990
A:Title: Analysis of variation in PR83, an erythrocytic merozoite vaccine candidate anti
A:Reference number: A44986; MUID:91101665
A:Accession: C44986
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: DNA
A:Residues: 1-622 <THO>
A:Cross-references: GB:M34554
C:Keywords: membrane protein; surface antigen
```

```
Query Match 100.0%; Score 105; DB 2; Length 622;
Best Local Similarity 100.0%; Pred. No. 3.7e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DDPKQYEQHLTDYEKIQEG 19
Db 348 DDPKQYEQHLTDYEKIQEG 366

RESULT 6
A44944 apical membrane antigen 1 precursor - Plasmodium fragile
C:Species: Plasmodium fragile
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-May-2000
C:Accession: A44944
R:Peterson, M.G.; Nguyen-Dinh, P.; Marshall, V.M.; Elliott, J.F.; Collins, W.E.; Ande
Mol. Biochem. Parasitol. 39, 279-284, 1990
A:Title: Apical membrane antigen of Plasmodium fragile.
A:Reference number: A44944; MUID:90205978
A:Accession: A44944
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-562 <PET>
A:Cross-references: GB:M29898

Query Match 77.1%; Score 81; DB 2; Length 562;
Best Local Similarity 73.7%; Pred. No. 0.00015;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 DDPKQYEQHLTDYEKIQEG 19
Db 293 DDPKQYEQHLTDYEKIQEG 311

RESULT 7
A39238 66K merozoite surface antigen precursor - Plasmodium knowlesi
C:Species: Plasmodium knowlesi
C:Date: 13-Sep-1991 #sequence_revision 13-Sep-1991 #text_change 07-Feb-1997
C:Accession: A39238
R:Waters, A.P.; Thomas, A.W.; Deans, J.A.; Mitchell, G.H.; Hudson, D.E.; Miller, L.H.
J. Biol. Chem. 265, 17974-17979, 1990
A:Title: A merozoite receptor protein from Plasmodium knowlesi is highly conserved an
A:Reference number: A39238; MUID:91009268
A:Accession: A39238
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-563 <WAT>
A:Cross-references: GB:J05631
C:Keywords: surface antigen

Query Match 75.2%; Score 79; DB 2; Length 563;
Best Local Similarity 73.7%; Pred. No. 0.00031;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 DDPKQYEQHLTDYEKIQEG 19
Db 293 DDPKQYEQHLTDYEKIQEG 311

RESULT 8
A44964 apical membrane antigen 1 - Plasmodium chabaudi adami
C:Species: Plasmodium chabaudi adami
C:Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 09-Sep-1997
C:Accession: A44964
R:Marshall, V.M.; Peterson, M.G.; Lew, A.M.; Kemp, D.J.
Mol. Biochem. Parasitol. 37, 281-283, 1989
A:Title: Structure of the apical membrane antigen I (AMA-1) of Plasmodium chabaudi.
A:Reference number: A44964; MUID:90114335
```

A:Accession: A44964
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-558 <MAR>
A:Cross-references: GB:M25248; NID:g160076; PID:g160077
C:Keywords: membrane protein; surface antigen

Query Match 68.6%; Score 72; DB 2; Length 558;
Best Local Similarity 68.4%; Pred. No. 0.0036;
Matches 13; Conservative 2; Mismatches 4; Indels 0;

Qy 1 DOPKQYEQHLTDYEKIKG 19
Db 293 DOPKQYKHLDTAKIRRG 311

RESULT 9

T31570
hypothetical protein Y105C5A.m - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T31570

R:McMurray, A.
submitted to the EMBL Data Library, September 1999

A:Reference number: Z21045
A:Accession: T31570

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-463 <WIL>

A:Cross-references: EMBL:AL117193; NID:e1549703; PIDN:CAB54993.1; CESP:Y105C5A.m

A:Experimental source: clone Y105C5A

C:Genetics:

A:Gene: CESP:Y105C5A.m

A:Introns: 15/1; 79/2; 114/3; 189/3; 237/1; 257/2; 401/3

Query Match 49.5%; Score 52; DB 2; Length 463;
Best Local Similarity 60.0%; Pred. No. 3.3;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 4 KQYEQHLTDYEKIKE 18
Db 64 KEYEMHLDTTKLEE 78

RESULT 10

T10738
hypothetical protein FbLate-2 - sea-island cotton

C:Species: Gossypium barbadense (sea-island cotton)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 17-Mar-2000

C:Accession: T10738

R:Rinehart, J.; Petersen, M.; John, M.E.

submitted to the EMBL Data Library, August 1995

A:Description: Tissue-specific and developmental regulation of cotton mRNA, FbLate-2: P1

A:Reference number: 217108

A:Accession: T10738

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-333 <RIN>

A:Cross-references: EMBL:U34401; NID:g1143223; PID:g1143224

A:Experimental source: cv. Sea Island

C:Genetics:

A:Gene: FbLate-2

C:Superfamily: proline-rich protein 3

Query Match 48.6%; Score 51; DB 2; Length 333;
Best Local Similarity 44.4%; Pred. No. 3.3;
Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DOPKQYEQHLTDYEKIKE 18
:::|::|:| :| | |

Db 211 EKPKEHEKHEVEYKPIKE 228

RESULT 11

C81263
probable integral membrane protein Cj1662 [imported] - Campylobacter jejuni (strain N

C:Species: Campylobacter jejuni

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Mar-2000

C:Accession: C81263

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kettle, J.M.; Churcher, C.; Basham, D.; Chil

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Ba

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals

A:Reference number: A81250; MUID:20150912

A:Accession: C81263

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-372 <PAR>

A:Cross-references: GB:AL139079; GB:AL111168; NID:g6968971; PIDN:CAB73649.1; PID:g696

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: Cj1662

Query Match 44.8%; Score 47; DB 2; Length 372;
Best Local Similarity 61.1%; Pred. No. 15;
Matches 11; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

Qy 3 PKQYEQHL--TDYEKIKE 18
| | | : | : | | | | |
Db 64 PKQDENFISNTYEKIKE 81

RESULT 12

I40644
botulinum neurotoxin type F nontoxic-nonhemagglutinin component - Clostridium botulinu

C:Species: Clostridium botulinum

C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 15-Oct-1999

C:Accession: I40644; S49002

R:East, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson,

FEMS Microbiol. Lett. 96, 225-230, 1992

A:Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.

A:Reference number: I40644

A:Accession: I40644

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1165 <RES>

A:Cross-references: EMBL:X71086; NID:g509740; PIDN:CAA50404.1; PID:g509741

R:Pochart, P.; Dore, J.; Lemann, F.; Goderel, I.; Rambaud, J.C.

FEMS Microbiol. Lett. 98, 225-228, 1992

A:Title: Interrelations between populations of methanogenic archaea and sulfate-reduc

A:Reference number: S49002

A:Accession: S49002

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1165 <POC>

A:Cross-references: EMBL:X71086; NID:g509740; PIDN:CAA50404.1; PID:g509741

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1993

C:Genetics:

A:Gene: ntnh

Query Match 44.8%; Score 47; DB 2; Length 1165;
Best Local Similarity 57.1%; Pred. No. 52;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DOPKQYEQHLTDYE 14
| | | : | : | | | | |

Db 257 DVPKVFEXKHKNDE 270

RESULT 13

E84486

hypothetical protein At2g07520 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: E84486
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Niernman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.W.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: E84486
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-222 <STO>
A:Cross-references: GB:AE002093; NID:g4895172; PIDN:AAD32760.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g07520
A:Map position: 2

Query Match 43.8%; Score 46; DB 2; Length 222;
Best Local Similarity 61.5%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 YEQHLDYKIKE 18
| | | | | | | |
DB 176 YAQHTDYQSK 188

RESULT 14
T35538
cytochrome-c oxidase (EC 1.9.3.1) chain II - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 04-Feb-2000
C:Accession: T35538
R:Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, March 1999
A:Reference number: Z21581
A:Accession: T35538
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-319 <SEE>
A:Cross-references: EMBL:AL049497; PIDN:CAB39883.1; GSPDB:GN00070; SCOEDB:SC6G10.29c
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: cox; SCOEDB:SC6G10.29c
C:Superfamily: cytochrome-c oxidase chain II; cytochrome-c oxidase chain II homology
C:Keywords: copper; electron transfer; membrane-associated complex; oxidoreductase; resp

Query Match 43.8%; Score 46; DB 2; Length 319;
Best Local Similarity 61.5%; Pred. No. 18;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 PKOYEQHLTDYK 15
| : | | | | | | |
DB 282 PERYEQLDLAK 294

RESULT 15
A70431
ribonucleotide reductase beta chain - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: A70431
R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666
A:Accession: A70431
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A:Residues: 1-696 <AQF>
A:Cross-references: GB:AE000743; NID:g2983875; PIDN:AAC07431.1; PID:g2983879; GB:AE00
A:Experimental source: strain VFS
C:Genetics:
A:Gene: nrdf

Query Match 43.8%; Score 46; DB 2; Length 696;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 KOYEQHLTDYK 15
| | | | | | | |
DB 72 KQETVLSDIK 83

Search completed: January 29, 2002, 10:26:45
Job time: 656 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:13:46 : Search time 80.65 Seconds
(without alignments)
8.638 Million cell updates/sec

Title: US-09-763-397A-18

Perfect score: 105

Sequence: 1 DQPKQYEQHLDYEIKKEG 19

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	100.0	622	1	ANAL_PLAF8
2	105	100.0	622	1	ANAL_PLAF8
3	105	100.0	622	1	ANAL_PLAF8
4	105	100.0	622	1	ANAL_PLAF8
5	105	100.0	622	1	ANAL_PLAF8
6	87	82.9	563	1	PK66_PLAKU
7	81	77.1	562	1	ANAL_PLAF8
8	72	68.6	588	1	ANAL_PLAF8
9	53	50.5	588	1	ANAL_PLAF8
10	45	42.9	332	1	ILVC_ARCFU
11	44	41.9	121	1	RL5_SOLME
12	44	41.9	256	1	Y448_METJA
13	44	41.9	457	1	FKB4_RABIT
14	44	41.9	459	1	FKB4_HUMAN
15	44	41.9	484	1	TTC2_HUMAN
16	44	41.9	986	1	GM13_RAT
17	44	41.9	1300	1	DYNA_NEUCR
18	43.5	41.4	70	1	Y416_UREPA
19	43.5	41.4	329	1	YOG1_CAEEL
20	43	41.0	332	1	ILVC_PYRAB
21	43	41.0	516	1	VLI_HPV05
22	43	41.0	614	1	GIDA_UREPA
23	43	41.0	1075	1	Y124_METJA
24	42	40.0	89	1	IN9A_HUMAN
25	42	40.0	89	1	IN9A_MOUSE
26	42	40.0	99	1	Y142_UREPA
27	42	40.0	359	1	RF1_UREPA
28	42	40.0	627	1	YERO_YEAST
29	42	40.0	693	1	YETI_SCHPO
30	42	40.0	777	1	UNC8_CAEEL
31	42	40.0	779	1	SRP_DROME
32	42	40.0	5255	1	BACA_BACLI
33	41	39.0	180	1	YE76_AQUAE

ALIGNMENTS

RESULT 1

ID	AMAL_PLAF8	STANDARD;	PRT;	622 AA.
AC	P50492;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).			
GN	AMA-1 OR PF83.			
OS	Plasmodium falciparum (isolate 7G8).			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
OX	NCBI_TaxID:57266;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-91101665; PubMed-2270110;			
RA	Thomas A.W., Waters A.P., Carr D.;			
RT	"Analysis of variation in Pf83, an erythrocytic merozoite vaccine candidate antigen of Plasmodium falciparum.";			
RL	Mol. Biochem. Parasitol. 42:285-287(1990).			
CC	-!- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.			
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	-!- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO PK66 FROM P.KNOWLES.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; M58548; AAA29721.1; .	
InterPro: IPR003298; Apmem_Ag1.	
Pfam; PF02430; AMA-1; 1.	
Malaria; Signal; Transmembrane; Antigen; Glycoprotein.	
FT SIGNAL	1 24
FT CHAIN	25 622
FT DOMAIN	25 546
FT TRANSMEM	547 567
FT DOMAIN	568 622
FT CARBOHYD	162 162
FT CARBOHYD	286 286
FT CARBOHYD	371 371
FT CARBOHYD	421 421
FT CARBOHYD	422 422
FT CARBOHYD	499 499
SQ SEQUENCE	622 AA; 71990 MW; IC9C8715D8E2915F CRC64;

Query Match 100.0%; Score 105; DB 1; Length 622;
Best Local Similarity 100.0%; Pred. No. 7e-08; 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;
QY 1 DQPKQYEQHLDYEIKKEG 19
|||||

```
Db 348 DQPKQYEQHLTDYEKKEG 366

RESULT 2
AMAL_PLAFC
ID AMAL_PLAFC STANDARD; PRT; 622 AA.
AC P50489;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
GN AMA-1 OR PF83.
OS Plasmodium falciparum (isolate Camp / Malaysia).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5835;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91101665; PubMed=2270110;
RA Thomas A.W., Waters A.P., Carr D.;
RT "Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate antigen of Plasmodium falciparum."
RL Mol. Biochem. Parasitol. 42:285-287(1990).
CC -1- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO PK66 FROM P.KNOWLES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M58545; AAA29718.1;
CC InterPro: IPR003298; Apmem_Agl.
CC Pfam; PF02430; AMA-1; 1.
CC Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
CC SIGNAL 1 24 POTENTIAL.
CC CHAIN 25 622 APICAL MEMBRANE ANTIGEN 1.
CC DOMAIN 25 546 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 547 567 POTENTIAL.
CC DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).
CC CARBOHYD 162 162 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 371 371 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 422 422 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 499 499 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CONFLICT 308 308 Q -> E (IN REF. 1; AAA29476).
CC CONFLICT 332 332 I -> N (IN REF. 1; AAA29476).
CC CONFLICT 407 407 Q -> H (IN REF. 1; AAA29476).
CC CONFLICT 439 439 H -> N (IN REF. 1; AAA29476).
CC CONFLICT 496 496 I -> M (IN REF. 1; AAA29476).
CC CONFLICT 503 503 N -> R (IN REF. 1; AAA29476).
CC SEQUENCE 622 AA; 71943 MW; 26CE8CF76D07C637 CRC64;

Query Match 100.0%; Score 105; DB 1; Length 622;
Best Local Similarity 100.0%; Pred. No. 7e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQYEQHLTDYEKKEG 19
| | | | | | | | | | | | | | | | | |
Db 348 DQPKQYEQHLTDYEKKEG 366

RESULT 3
AMAL_PLAFC
ID AMAL_PLAFC STANDARD; PRT; 622 AA.
AC P22621;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
GN AMA-1 OR PF83.
OS Plasmodium falciparum (isolate FC27 / Papua New Guinea).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5835;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91101665; PubMed=2270110;
RA Thomas A.W., Waters A.P., Carr D.;
RT "Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate antigen of Plasmodium falciparum."
RL Mol. Biochem. Parasitol. 42:285-287(1990).
CC -1- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO PK66 FROM P.KNOWLES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M58545; AAA29718.1;
CC InterPro: IPR003298; Apmem_Agl.
CC Pfam; PF02430; AMA-1; 1.
CC Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
CC SIGNAL 1 24 POTENTIAL.
CC CHAIN 25 622 APICAL MEMBRANE ANTIGEN 1.
CC DOMAIN 25 546 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 547 567 POTENTIAL.
CC DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).
CC CARBOHYD 162 162 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 371 371 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 422 422 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 499 499 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CONFLICT 308 308 Q -> E (IN REF. 1; AAA29476).
CC CONFLICT 332 332 I -> N (IN REF. 1; AAA29476).
CC CONFLICT 407 407 Q -> H (IN REF. 1; AAA29476).
CC CONFLICT 439 439 H -> N (IN REF. 1; AAA29476).
CC CONFLICT 496 496 I -> M (IN REF. 1; AAA29476).
CC CONFLICT 503 503 N -> R (IN REF. 1; AAA29476).
CC SEQUENCE 622 AA; 71943 MW; 26CE8CF76D07C637 CRC64;

Query Match 100.0%; Score 105; DB 1; Length 622;
Best Local Similarity 100.0%; Pred. No. 7e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQYEQHLTDYEKKEG 19
| | | | | | | | | | | | | | | | | |
Db 348 DQPKQYEQHLTDYEKKEG 366
```

```
OX NCBI_TaxID=5837;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89384584; PubMed=2701947;
RA Peterson M.G., Marshall V.M., Smythe J.A., Crewther P.E., Lew A., Silva A., Anders R.F., Kemp D.J.;
RT "Integral membrane protein located in the apical complex of Plasmodium falciparum."
RL Mol. Cell. Biol. 9:3151-3154(1989).
CC -1- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO PK66 FROM P.KNOWLES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M27133; AAA29475.1;
CC EMBL; M27957; AAA29476.1;
CC EMBL; A08267; CAA00764.1;
CC PIR; A32499; A32499.
CC InterPro: IPR003298; Apmem_Agl.
CC Pfam; PF02430; AMA-1; 1.
CC Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
CC SIGNAL 1 24 POTENTIAL.
CC CHAIN 25 622 APICAL MEMBRANE ANTIGEN 1.
CC DOMAIN 25 546 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 547 567 POTENTIAL.
CC DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).
CC CARBOHYD 162 162 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 371 371 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 422 422 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 499 499 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CONFLICT 308 308 Q -> E (IN REF. 1; AAA29476).
CC CONFLICT 332 332 I -> N (IN REF. 1; AAA29476).
CC CONFLICT 407 407 Q -> H (IN REF. 1; AAA29476).
CC CONFLICT 439 439 H -> N (IN REF. 1; AAA29476).
CC CONFLICT 496 496 I -> M (IN REF. 1; AAA29476).
CC CONFLICT 503 503 N -> R (IN REF. 1; AAA29476).
CC SEQUENCE 622 AA; 72009 MW; 7D41335E249FA18F CRC64;

Query Match 100.0%; Score 105; DB 1; Length 622;
Best Local Similarity 100.0%; Pred. No. 7e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQYEQHLTDYEKKEG 19
| | | | | | | | | | | | | | | | | |
Db 348 DQPKQYEQHLTDYEKKEG 366

RESULT 4
AMAL_PLAFC
ID AMAL_PLAFC STANDARD; PRT; 622 AA.
AC P50490;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
GN AMA-1 OR PF83.
OS Plasmodium falciparum (isolate FCR-3 / Gambia).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5838;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91101665; PubMed=2270110;
```


RA Thomas A.W., Waters A.P., Carr D.;
RT "Analysis of variation in PF83, an erythrocytic merozoite vaccine
RL candidate antigen of Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 42:285-287(1990).
CC -1- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO
CC PK66 FROM P.KNOWLES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M58546; AAA29719.1; -;
DR InterPro: IPR003298; Apmem_Agl.
DR Pfam: PF02430; AMA-1; 1.
KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
FT SIGNAL 1 24
FT CHAIN 25 622
FT DOMAIN 25 546
FT TRANSMEM 547 567
FT DOMAIN 568 622
FT CARBOHYD 162 162
FT CARBOHYD 286 286
FT CARBOHYD 371 371
FT CARBOHYD 421 421
FT CARBOHYD 422 422
FT CARBOHYD 499 499
FT CARBOHYD 622 AA; 089336BE0464695C CRC64;
SQ SEQUENCE 622 AA; 71967 MW; 089336BE0464695C CRC64;

Query Match 100.0%; Score 105; DB 1; Length 622;
Best Local Similarity 100.0%; Pred. No. 7e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQYEQHLTDYKIKEG 19
Db 348 DQPKQYEQHLTDYKIKEG 366

RESULT 5
AMAL_PLAFH STANDARD; PRT; 622 AA.
AC P50491;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
GN AMA-1 OR PF83.
OS Plasmodium falciparum (isolate thtn / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=70151;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91101665; PubMed=2270110;
RA Thomas A.W., Waters A.P., Carr D.;
RT "Analysis of variation in PF83, an erythrocytic merozoite vaccine
RL candidate antigen of Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 42:285-287(1990).
CC -1- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO
CC PK66 FROM P.KNOWLES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M58547; AAA29720.1; -;
DR InterPro: IPR003298; Apmem_Agl.
DR Pfam: PF02430; AMA-1; 1.
KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
FT SIGNAL 1 24
FT CHAIN 25 622
FT DOMAIN 25 546
FT TRANSMEM 547 567
FT DOMAIN 568 622
FT CARBOHYD 286 286
FT CARBOHYD 371 371
FT CARBOHYD 421 421
FT CARBOHYD 422 422
FT CARBOHYD 499 499
FT CARBOHYD 622 AA; 71989 MW; 1FDFA53593C94CC5 CRC64;
SQ SEQUENCE 622 AA; 71989 MW; 1FDFA53593C94CC5 CRC64;

Query Match 100.0%; Score 105; DB 1; Length 622;
Best Local Similarity 100.0%; Pred. No. 7e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQYEQHLTDYKIKEG 19
Db 348 DQPKQYEQHLTDYKIKEG 366

RESULT 6
PK66_PLAKU STANDARD; PRT; 563 AA.
AC P21303;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE MEROZOITE RECEPTOR PK66 PRECURSOR (66 KDA PROTECTIVE MINOR SURFACE
DE ANTIGEN).
GN PK66.
OS Plasmodium knowlesi (strain nuri).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5852;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LINE W1;
RX MEDLINE=91009268; PubMed=2211675;
RA Waters A.P., Thomas A.W., Deans J.A., Mitchell G.H., Hudson D.E.,
RA Miller L.H., McCutchan T.F., Cohen S.;
RT "A merozoite receptor protein from Plasmodium knowlesi is highly
RT conserved and distributed throughout Plasmodium.";
RL J. Biol. Chem. 265:17974-17979(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=NURI;
RX MEDLINE=91187058; PubMed=2011149;
RA Waters A.P., Thomas A.W., Mitchell G.H., McCutchan T.F.;
RT "Intra-generic conservation and limited inter-strain variation in a
RT protective minor surface antigen of Plasmodium knowlesi merozoites.";
RL Mol. Biochem. Parasitol. 44:141-144(1991).
CC -1- FUNCTION: MEROZOITE RECEPTOR PK66 IS A SURFACE ANTIGEN INVOLVED
CC IN PARASITE INVASION OF ERYTHROCYTES.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CONCENTRATED AT THE
CC APICAL END PRIOR TO RUPTURE, FOLLOWING WHICH IT CAN DISTRIBUTE
CC ITSELF ENTIRELY ACROSS THE SURFACE OF THE FREE MEROZOITE. DURING
CC INVASION PK66 IS EXCLUDED FROM THE ERYTHROCYTE AT, AND BEHIND, THE
CC INVASION INTERFACE.
CC -1- MISCELLANEOUS: PK66 EXPRESSED IN MATURE SCHIZONTS IS RAPIDLY
CC PROCESSED AS THE SCHIZONT RUPTURES, YIELDING A 42/44 KDA DOUBLET
CC ASSOCIATED WITH THE SURFACE MEROZOITE.
CC -1- SIMILARITY: STRONG TO AMA-1/PF83 OF P.FALCIPARUM, P.CHABAUDI, AND
CC P.FRAGILE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M58317; AAA63444.1; -;
 DR EMBL; M61097; AAA29728.1; -;
 DR InterPro; IPR003298; Apmem_Ag1.
 DR Pfam; PF02430; AMA-1; 1.
 KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
 FT SIGNAL 1 13 POTENTIAL.
 FT CHAIN 14 563 MEROZOITE RECEPTOR PK66.
 FT DOMAIN 14 487 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 488 508 POTENTIAL.
 FT DOMAIN 509 563 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 107 107 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 228 228 N -> K (IN STRAIN LINE W1).
 SQ SEQUENCE 563 AA; 64680 MW; 9EF06202644A5CCF CRC64;

Query Match 82.9%; Score 87; DB 1; Length 563;
 Best Local Similarity 78.9%; Pred. No. 2.6e-05;
 Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 DQPKQYEQHLTDYKIKEG 19
 DB 293 DQPROVEEELTDYKIQEG 311
 ||| ||| ||| ||| ||| |||

RESULT 7
 AMAL_PLAIFR STANDARD; PRT; 562 AA.
 AC P22622;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
 GN AMA-1 OR AG352.
 OS Plasmodium fragile.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5857;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90205978; PubMed=2181309;
 RA Peterson M.G., Nguyen-Dinh P., Marshall V.M., Elliott J.F.,
 RA Collins W.B., Anders R.F., Kemp D.J.;
 RT "Apical membrane antigen of Plasmodium fragile."
 RL Mol. Biochem. Parasitol. 39:279-284(1990).
 CC -!- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FALCIPARUM, AND
 CC TO PK66 FROM P.KNOWLES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M29898; AAA29474.1; -;
 DR PIR; A44944; A44944.
 DR InterPro; IPR003298; Apmem_Ag1.
 DR Pfam; PF02430; AMA-1; 1.
 KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.

FT SIGNAL 1 21 POTENTIAL.

FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 562 APICAL MEMBRANE ANTIGEN 1.
 FT DOMAIN 22 484 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 485 507 POTENTIAL.
 FT DOMAIN 508 562 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 562 AA; 64498 MW; 9EAB72D437EA7164 CRC64;

Query Match 77.1%; Score 81; DB 1; Length 562;
 Best Local Similarity 73.7%; Pred. No. 0.00019;
 Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 DQPKQYEQHLTDYKIKEG 19
 DB 293 DQPROVEEELTDYKIQEG 311
 ||| ||| ||| ||| ||| |||

RESULT 8

AMAL_PLAICH STANDARD; PRT; 558 AA.
 ID AMA1_PLACH
 AC P16445;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
 GN AMA-1.
 OS Plasmodium chabaudi.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5825;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DS;
 RX MEDLINE=90114335; PubMed=2608101;
 RA Marshall V.M., Peterson M.G., Lew A.M., Kemp D.J.;
 RT "Structure of the apical membrane antigen 1 (AMA-1) of Plasmodium
 RT chabaudi."
 RL Mol. Biochem. Parasitol. 37:281-284(1989).
 RN [2]
 RP SEQUENCE OF 38-377 FROM N.A.
 RC STRAIN=DK;
 RX MEDLINE=96333375; PubMed=8757869;
 RA Crewe P.E., Matthew M.L., Flegg R.H., Anders R.F.;
 RT "Protective immune responses to apical membrane antigen 1 of
 RT Plasmodium chabaudi involve recognition of strain-specific
 RT epitopes."
 RL Infect. Immun. 64:3310-3317(1996).
 CC -!- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- SIMILARITY: STRONG TO AMA-1/PF83 FROM P.FALCIPARUM, P.FRAGILE AND
 CC TO PK66 FROM P.KNOWLES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M25248; AAA90929.1; -;
 DR EMBL; A08270; CAA00765.1; -;
 DR EMBL; U49745; AAB36511.1; -;
 DR PIR; A44964; A44964.
 DR InterPro; IPR003298; Apmem_Ag1.
 DR Pfam; PF02430; AMA-1; 1.
 KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 558 APICAL MEMBRANE ANTIGEN 1.

FT DOMAIN 22 480 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 481 503 POTENTIAL.
FT DOMAIN 504 558 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 20 20 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 249 249 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 412 412 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 558 AA; 63973 MW; 9773F3E6A439A972 CRC64;

Query Match 68.6%; Score 72; DB 1; Length 558;
Best Local Similarity 68.4%; Pred. No. 0.0038;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 DQPKYEQHLTDYEKIK 19
DB 293 DQPKYEQHLTDYEKIRRG 311
|||||:|:| | | |

RESULT 9
DYR3_HUMAN STANDARD; PRT; 588 AA.
AC O43781;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DUAL-SPECIFICITY TYROSINE-PHOSPHORYLATION REGULATED KINASE 3
DE (EC 2.7.1.-).
GN DYRK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]

RN SEQUENCE FROM N.A.
RP TISSUE=Fetal brain;
RX MEDLINE=98421512; PubMed=9748265;
RA Becker W., Weber Y., Wetzel K., Eimbter K., Tejedor F.J.,
RA Joost H.-G.;
RT "Sequence characteristics, subcellular localization, and substrate
RT specificity of DYRK-related kinases, a novel family of dual
RT specificity protein kinases.";
RN J. Biol. Chem. 273:25893-25902(1998).
[2]

RN REVISIONS TO N-TERMINAL.
RA Becker W.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RX MEDLINE=99063810; PubMed=9845759;
RA Xia J., Tang X., Ruan Q., Pan Q., Liu C., Xie W., Deng H.;
RT "Molecular cloning and characterization of novel protein kinase gene
RT DYRK3.";
RN Zhonghua Yi Xue Yi Chuan Xue Za Zhi 15:327-332(1998).

CC -1- FUNCTION: IN VITRO: CAN PHOSPHORYLATE HISTONES H3 AND H2B ON SER
CC AND THR RESIDUES. MAY BE INVOLVED IN THE REGULATION OF CELLULAR
CC GROWTH AND/OR DEVELOPMENT.
CC -1- PPM: AUTOPHOSPHORYLATED ON TYR RESIDUES.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MNB/DYRK SUBFAMILY.
[1]

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

DR EMBL; Y12735; CAA73266.2; -

DR HSP: Q16539; 1A9U.
DR MIM: 603497;
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_kin_actsite.
DR Pfam: PF00069; pkinase; 4.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding; Phosphorylation.
KW DOMAIN 209 522 PROTEIN KINASE.
FT NP_BIND 215 223 ATP (BY SIMILARITY).
FT BINDING 238 238 ATP (BY SIMILARITY).
FT ACT_SITE 335 335 BY SIMILARITY.
SQ SEQUENCE 588 AA; 65812 MW; D43D897D34C29D9E CRC64;

Query Match 50.5%; Score 53; DB 1; Length 588;
Best Local Similarity 64.3%; Pred. No. 2.3;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 KQYEQHLTDYEKIK 17
DB 148 KQYKHLTAYEKL 161
|||: ||| |||:

RESULT 10
ILVC_ARCFU STANDARD; PRT; 332 AA.
ID ILVC_ARCFU
AC O28294;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE KETOL-ACID REDUCTOISOMERASE (EC 1.1.1.86) (ACETOHYDROXY-ACID
DE ISOMEROREDUCTASE) (ALPHA-KETO-BETA-HYDROXYLACIL REDUCTOISOMERASE).
GN ILVC OR AF1985.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
[1]

RN SEQUENCE FROM N.A.
RP STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spraggins T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;

RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -1- CATALYTIC ACTIVITY: (R)-2,3-DIHYDROXY-3-METHYLBUTANONATE + NADP(+)
CC = (S)-2-HYDROXY-2-METHYL-3-OXOBUTANONATE + NADPH.
CC -1- PATHWAY: VALINE AND ISOLEUCINE BIOSYNTHESIS; SECOND STEP.
CC -1- SIMILARITY: BELONGS TO THE KETOL-ACID REDUCTOISOMERASE FAMILY.
[1]

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

DR EMBL; AE000966; AAB89269.1; -

DR TIGR; AF1985; -
DR InterPro; IPR000506; Acetylcholinesterase; Methanococcales; Methanococcaceae;
DR Pfam; PF01450; ILVC; 1.
KW Oxidoreductase; Branched-chain amino acid biosynthesis; NADP;
KW Complete proteome.
FT ACT_SITE 109 POTENTIAL.
SQ SEQUENCE 332 AA; 37212 MW; 598EF6AB3E2B000F CRC64;

Query Match 42.9%; Score 45; DB 1; Length 332;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 9; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY 2 QPKOYEQHLTDYEKIKG 19
||| :|| :|||
Db 87 QPAVYREHIQD--KLKGG 102

RESULT 11
RL5_SOLME STANDARD; PRT; 121 AA.
AC P93779;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 13-DEC-1998 (Rel. 37, Last annotation update)
DE 60S RIBOSOMAL PROTEIN L5 (FRAGMENT).
GN RPL5 OR TM003.
OS Solanum melongena (Eggplant) (Aubergine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4111;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cotyledon;
RA Momiyama T., Kayano T., Takaiwa F., Takayanagi K.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN BINDS 5S RNA (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB001583; BAA19415.1; -
DR DR
KW Ribosomal protein.
FT NON_TER 1
SQ SEQUENCE 121 AA; 13979 MW; D0CBF76A71FF02B CRC64;

Query Match 41.9%; Score 44; DB 1; Length 121;
Best Local Similarity 40.0%; Pred. No. 9.3;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 DPKOYEQHLTDYEK 15
||| :|| :|||
Db 32 DEPEKQSHFSSYIK 46

RESULT 12
Y448_METJA STANDARD; PRT; 256 AA.
ID Y448_METJA
AC C57850;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHEICAL PROTEIN MJ0448.
GN MJ0448.
OS Methanococcus jannaschii.
OC Methanococcus jannaschii.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Lebeau M.-C., Massol N., Herrick J., Faber L.E., Renoir J.-M.,
RA Radanyi C., Baulieu E.-E.;
RT "p59, an hsp 90-binding protein. Cloning and sequencing of its cDNA
RT and preparation of a peptide-directed polyclonal antibody.";
RL J. Biol. Chem. 267:4281-4284(1992).
RN [2]
RP DOMAINS.
RX MEDLINE=92335279; PubMed=1631118;
RA Callebaut I., Renoir J.-M., Lebeau M.-C., Massol N., Burry A.,
RA Baulieu E.-E., Mornon J.-P.;
RT "An immunophilin that binds M(r) 90,000 heat shock protein: main

structural features of a mammalian p59 protein.":
Proc. Natl. Acad. Sci. U.S.A. 89:6270-6274(1992).
[3]
SEQUENCE OF 1-25
STRAIN-NEW ZEALAND WHITE; TISSUE=Liver;
RX MEDLINE=96154240; PubMed=8579355;
RA Deshpande K.L., Seubert P.H., Tillman D.M., Farkas W.R., Katze J.R.;
RT "Cloning and characterization of cDNA encoding the rabbit
RNA-guanine transglycosylase 60-kilodalton subunit.";
Arch. Biochem. Biophys. 326:1-7(1996).
[4]
STRUCTURE BY NMR OF 1-148.
RX MEDLINE=96374215; PubMed=8780506;
RA Craescu C.T., Rouviere N., Popescu A., Cerpolini E., Lebeau M.-C.,
RA Baulieu E.-E., Mispelter J.;
RT "Three-dimensional structure of the immunophilin-like domain of
FKBP59 in solution";
Biochemistry 35:11045-11052(1996).
CC -!- FUNCTION: BINDS TO HSP90 AND IS ASSOCIATED WITH IT, INCLUDING
CC WHEN HSP90 PARTICIPATES IN HETERO-OLIGOMERIC COMPLEXES OF
CC UNTRANSFORMED MAMMALIAN STEROID RECEPTORS THAT SEDIMENT AT 8-10 S.
CC P59 MAY HAVE A ROTAMASE ACTIVITY WHICH MAY APPLY TO BOUND HSP90
CC AND THUS BE IMPLIED IN THE INTRACELLULAR TRAFFICKING OF HETERO-
CC OLIGOMERIC FORMS OF STEROID HORMONE RECEPTORS.
CC -!- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
CC PEPTIDE BONDS IN OLIGOPEPTIDES.
CC -!- SUBUNIT: INTERACTS WITH PEROXISOMAL PHYTANYL-COA ALPHA-
CC HYDROXYLASE.
CC -!- SUBCELLULAR LOCATION: NUCLEAR AND CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE FKBP-TYPE PPIASE FAMILY. CONTAINS 2
CC FKBP-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 3 TPR REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M84474; AAA31438.1; .
DR EMBL; M84988; AAA31439.1; .
DR PIR; A42386; A42386.
DR PDB; 1ROT; 07-DEC-96.
DR PDB; 1ROU; 07-DEC-96.
DR InterPro; IPR001179; FKBP_PPIase.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00254; FKBP; 2.
DR Pfam; PF00515; TPR; 3.
DR PROSITE; PS00453; FKBP_PPIASE_1; 1.
DR PROSITE; PS00454; FKBP_PPIASE_2; 2.
DR PROSITE; PS00509; FKBP_PPIASE_3; 2.
KW Isomerase; Rotamase; TPR repeat; Repeat; Nuclear protein;
KW 3D-structure.
FT INIT_MET 0
FT DOMAIN 49 137 PPIASE, FKBP-TYPE 1.
FT DOMAIN 166 252 PPIASE, FKBP-TYPE 2.
FT REPEAT 269 302 TPR 1.
FT REPEAT 318 351 TPR 2.
FT REPEAT 353 385 TPR 3.
FT CONFLICT 14 21 S -> H (IN REF. 3).
FT CONFLICT 20 21 EG -> FI (IN REF. 3).
FT CONFLICT 25 25 S -> T (IN REF. 3).
SQ SEQUENCE 457 AA; 51344 MW; EC58CC4BCF66A44A CRC64;

Query Match 41.9%; Score 44; DB 1; Length 457;
Best Local Similarity 57.18; Pred. No. 35;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Oy 5 QYEHLTDYEKIKE 18
:||| :|||

Db 243 KYEVLKSFKEAKE 256
RESULT 14
FKB4_HUMAN
ID FKB4_HUMAN STANDARD; PRT; 459 AA.
AC Q02790;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE P59 PROTEIN (HSP BINDING IMMUNOPHILIN) (HBI) (POSSIBLE PEPTIDYL-PROLYL
DE CIS-TRANS ISOMERASE) (EC 5.2.1.8) (PPIASE) (ROTAMASE) (FKBP52 PROTEIN)
DE (52 KDA FK506 BINDING PROTEIN) (P52) (FKBP59) (HSP56).
GN FKBP4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92147620; PubMed=13711107;
RA Yem A.W., Tonaselli A.G., Heinrikson R.L., Zurcher-Neely H.,
RA Ruff V.A., Johnson R.A., Deibel M.R.;
RT "The hsp56 component of steroid receptor complexes binds to
RT immobilized FK506 and shows homology to FKBP-12 and FKBP-13.";
J. Biol. Chem. 267:2868-2871(1992).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93066366; PubMed=1279700;
RA Peattie D.A., Harding M.W., Fleming M.A., Decenzo M.T.,
RA Lippke J.A., Livingston D.J., Benasutti M.;
RT "Expression and characterization of human FKBP52, an immunophilin
RT that associates with the 90-kDa heat shock protein and is a component
RT of steroid receptor complexes";
Proc. Natl. Acad. Sci. U.S.A. 89:10974-10978(1992).
RL -!- FUNCTION: BINDS TO HSP90 AND IS ASSOCIATED WITH IT, INCLUDING
CC WHEN HSP90 PARTICIPATES IN HETERO-OLIGOMERIC COMPLEXES OF
CC UNTRANSFORMED MAMMALIAN STEROID RECEPTORS THAT SEDIMENT AT 8-10 S.
CC P59 MAY HAVE A ROTAMASE ACTIVITY WHICH MAY APPLY TO BOUND HSP90
CC AND THUS BE IMPLIED IN THE INTRACELLULAR TRAFFICKING OF HETERO-
CC OLIGOMERIC FORMS OF STEROID HORMONE RECEPTORS.
CC -!- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
CC PEPTIDE BONDS IN OLIGOPEPTIDES.
CC -!- SUBUNIT: INTERACTS WITH PEROXISOMAL PHYTANYL-COA ALPHA-
CC HYDROXYLASE.
CC -!- SUBCELLULAR LOCATION: NUCLEAR AND CYTOPLASMIC.
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED.
CC -!- SIMILARITY: BELONGS TO THE FKBP-TYPE PPIASE FAMILY. CONTAINS 2
CC FKBP-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 3 TPR REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M88279; AAA36111.1; .
DR PIR; A46372; A46372.
DR HSSP; P27124; 1ROU.
DR MM; 600611; .
DR InterPro; IPR001179; FKBP_PPIase.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00254; FKBP; 2.
DR Pfam; PF00515; TPR; 3.
DR PROSITE; PS00453; FKBP_PPIASE_1; 1.
DR PROSITE; PS00454; FKBP_PPIASE_2; 2.
DR PROSITE; PS00509; FKBP_PPIASE_3; 2.
KW Isomerase; Rotamase; Repeat; TPR repeat; Nuclear protein.
FT DOMAIN 50 138 PPIASE, FKBP-TYPE 1.
FT DOMAIN 167 253 PPIASE, FKBP-TYPE 2.

FT REPEAT 270 303 TPR 1.
FT REPEAT 319 352 TPR 2.
FT REPEAT 354 386 TPR 3.
SQ SEQUENCE 459 AA; 51804 MW; 6A498105418D9435 CRC64;

Query Match 41.9%; Score 44; DB 1; Length 459;
Best Local Similarity 57.1%; Pred. No. 36;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 QYEQHLTDYEKIKE 18
: || || : || ||
Db 244 KYELHLKSEKAKE 257

RESULT 15
TTC2_HUMAN
ID TTC2_HUMAN STANDARD; PRT; 484 AA.
AC Q99615;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TETRATRICPEPTIDE REPEAT PROTEIN 2 (TPR REPEAT PROTEIN 2).
GN TTC2 OR TPR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96433003; PubMed=8836031;
RA Murthy A.E., Bernards A., Church D., Wasmuth J., Gusella J.F.;
RT "Identification and characterization of two novel tetratricopeptide
repeat-containing genes."
RL DNA Cell Biol. 15:727-735(1996).
CC -!- SUBUNIT: INTERACTS WITH THE GAP DOMAIN OF NF1.
CC -!- SIMILARITY: CONTAINS 9 TPR REPEATS.
CC -!- SIMILARITY: CONTAINS 1 J DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U46571; AAB36872.1; -.
DR HSP; P08622; 1XBL.
DR MIM; 601964; -.
DR InterPro; IPR001623; DnaJ_N.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00226; DnaJ; 1.
DR Pfam; PF00515; TPR; 8.
DR SMART; SM00271; DnaJ; 1.
DR SMART; SM00028; TPR; 7.
DR PROSITE; PS00636; DnaJ_1; FALSE_NEG.
DR PROSITE; PS50076; DnaJ_2; 1.
KW Repeat; TPR repeat; Chaperone.
FT REPEAT 18 51 TPR 1.
FT REPEAT 53 85 TPR 2.
FT REPEAT 86 119 TPR 3.
FT REPEAT 132 165 TPR 4.
FT REPEAT 167 199 TPR 5.
FT REPEAT 200 233 TPR 6.
FT REPEAT 246 279 TPR 7.
FT REPEAT 284 317 TPR 8.
FT REPEAT 318 351 TPR 9.
FT DOMAIN 371 441 J-DOMAIN.
SQ SEQUENCE 484 AA; 55479 MW; 4147EBDB0CES0DB2 CRC64;

Query Match 41.9%; Score 44; DB 1; Length 484;

Best Local Similarity 53.8%; Pred. No. 38;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 4 KOYEQLHTDYEKI 16
: || || : || ||
Db 332 EQYEEAVRDYEV 344

Search completed: January 29, 2002, 11:13:47
Job time: 823 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:12:14 ; Search time 285.36 Seconds
(without alignments)
9.739 Million cell updates/sec

Title: US-09-763-397A-18

Perfect score: 105

Sequence: 1 DQPKQYEQLTDYKKEG 19

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	105	100.0	402	5 Q9BIM8	Q9BIM8 plasmodium
2	105	100.0	437	5 Q9BHS1	Q9BHS1 plasmodium
3	105	100.0	437	5 Q9BHS0	Q9BHS0 plasmodium
4	105	100.0	437	5 Q9BHR9	Q9BHR9 plasmodium
5	105	100.0	437	5 Q9BHR8	Q9BHR8 plasmodium
6	105	100.0	437	5 Q9BHR7	Q9BHR7 plasmodium
7	105	100.0	437	5 Q9BHR6	Q9BHR6 plasmodium
8	105	100.0	437	5 Q9BHR5	Q9BHR5 plasmodium
9	105	100.0	437	5 Q9BHR4	Q9BHR4 plasmodium
10	105	100.0	437	5 Q9BHR3	Q9BHR3 plasmodium
11	105	100.0	437	5 Q9BHR2	Q9BHR2 plasmodium
12	105	100.0	437	5 Q9BHR1	Q9BHR1 plasmodium
13	105	100.0	437	5 Q9BHR0	Q9BHR0 plasmodium
14	105	100.0	437	5 Q9BHQ9	Q9BHQ9 plasmodium
15	105	100.0	437	5 Q9BHQ8	Q9BHQ8 plasmodium
16	105	100.0	437	5 Q9BHQ7	Q9BHQ7 plasmodium
17	105	100.0	437	5 Q9BHQ6	Q9BHQ6 plasmodium
18	105	100.0	437	5 Q9BHQ5	Q9BHQ5 plasmodium
19	105	100.0	437	5 Q9BHQ4	Q9BHQ4 plasmodium

20	105	100.0	437	5 Q9BHQ3	Q9BHQ3 plasmodium
21	105	100.0	437	5 Q9BHQ2	Q9BHQ2 plasmodium
22	105	100.0	437	5 Q9BHQ1	Q9BHQ1 plasmodium
23	105	100.0	437	5 Q9BHQ0	Q9BHQ0 plasmodium
24	105	100.0	437	5 Q9BHP9	Q9BHP9 plasmodium
25	105	100.0	437	5 Q9BHP8	Q9BHP8 plasmodium
26	105	100.0	437	5 Q9BHP7	Q9BHP7 plasmodium
27	105	100.0	437	5 Q9BHP6	Q9BHP6 plasmodium
28	105	100.0	437	5 Q9BHP5	Q9BHP5 plasmodium
29	105	100.0	437	5 Q9BHP4	Q9BHP4 plasmodium
30	105	100.0	437	5 Q9BHP3	Q9BHP3 plasmodium
31	105	100.0	437	5 Q9BHP2	Q9BHP2 plasmodium
32	105	100.0	437	5 Q9BHP1	Q9BHP1 plasmodium
33	105	100.0	437	5 Q9BHP0	Q9BHP0 plasmodium
34	105	100.0	437	5 Q9BHN9	Q9BHN9 plasmodium
35	105	100.0	437	5 Q9BHN8	Q9BHN8 plasmodium
36	105	100.0	437	5 Q9BHN7	Q9BHN7 plasmodium
37	105	100.0	437	5 Q9BHN6	Q9BHN6 plasmodium
38	105	100.0	437	5 Q9BHN5	Q9BHN5 plasmodium
39	105	100.0	437	5 Q9BHN4	Q9BHN4 plasmodium
40	105	100.0	437	5 Q9BHN3	Q9BHN3 plasmodium
41	105	100.0	437	5 Q9BHN2	Q9BHN2 plasmodium
42	105	100.0	437	5 Q9BHB6	Q9BHB6 plasmodium
43	105	100.0	437	5 Q9BHB4	Q9BHB4 plasmodium
44	105	100.0	437	5 Q9BHB7	Q9BHB7 plasmodium
45	105	100.0	437	5 Q9BHB26	Q9BHB26 plasmodium

ALIGNMENTS

RESULT 1
Q9BIM8 PRELIMINARY; PRT; 402 AA.
AC Q9BIM8;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE APICAL MEROZOITE ANTIGEN-1 (FRAGMENT).
GN AMA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN-FCB-1;
RA Goel V.K., Chishti A.H., Oh S.S.;
RT "Plasmodium falciparum apical membrane antigen-1 (AMA-1) gene, partial
CDNA sequence."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF352829; AAK26116.1;
KW Merozoite.
FT NON_TER 1
FT NON_TER 402
SQ SEQUENCE 402 AA; 46227 MW; 06DAE211FD596275 CRC64;

Query Match 100.0%; Score 105; DB 5; Length 402;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQPKQYEQLTDYKKEG 19
|||||
Db 212 DQPKQYEQLTDYKKEG 230

RESULT 2
Q9BHS1 PRELIMINARY; PRT; 437 AA.
AC Q9BHS1;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)

```
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=NIGERIAN 002;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical
  Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ408300; CAC34741.1; -.
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 50218 MW; C043561972E8D49F CRC64;

Query Match 100.0%; Score 105; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQYEQLTDYEKKEG 19
Db 202 DQPKQYEQLTDYEKKEG 220
|||||

RESULT 3
Q9BHS0 ID Q9BHS0 PRELIMINARY; PRT; 437 AA.
AC Q9BHS0;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=NIGERIAN 006;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical
  Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ408302; CAC34743.1; -.
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 50137 MW; B4D79C665B529AFA CRC64;

Query Match 100.0%; Score 105; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQYEQLTDYEKKEG 19
Db 202 DQPKQYEQLTDYEKKEG 220
|||||

RESULT 4
Q9BHR9 ID Q9BHR9 PRELIMINARY; PRT; 437 AA.
AC Q9BHR9;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
```

```
RN SEQUENCE FROM N.A.
RC STRAIN=NIGERIAN 015;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical
  Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ408303; CAC34744.1; -.
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 50243 MW; 28D9E9E5165A39EE CRC64;

Query Match 100.0%; Score 105; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQYEQLTDYEKKEG 19
Db 202 DQPKQYEQLTDYEKKEG 220
|||||

RESULT 5
Q9BHR8 ID Q9BHR8 PRELIMINARY; PRT; 437 AA.
AC Q9BHR8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=NIGERIAN 029;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical
  Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ408306; CAC34747.1; -.
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 50247 MW; 8C6C4F7A87139B70 CRC64;

Query Match 100.0%; Score 105; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQYEQLTDYEKKEG 19
Db 202 DQPKQYEQLTDYEKKEG 220
|||||

RESULT 6
Q9BHR7 ID Q9BHR7 PRELIMINARY; PRT; 437 AA.
AC Q9BHR7;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=NIGERIAN 030;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical
```


Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQYEQHLYDYKKEG 19
|||||
Db 202 DQPKQYEQHLYDYKKEG 220

RESULT 11

Q9BHR2 ID Q9BHR2 PRELIMINARY; PRT; 437 AA.

AC Q9BHR2; 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1.

OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=NIGERIAN 039;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical
RT Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ408312; CAC34753.1; -.
FT NON_TER 1
FT NON_TER 437 437
SQ SEQUENCE 437 AA; 50356 MW; 734A74080D1BAC7F CRC64;

Query Match 100.0%; Score 105; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQYEQHLYDYKKEG 19
|||||
Db 202 DQPKQYEQHLYDYKKEG 220

RESULT 12

Q9BHR1 ID Q9BHR1 PRELIMINARY; PRT; 437 AA.

AC Q9BHR1; 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1.

OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=NIGERIAN 044;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical
RT Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ408314; CAC34755.1; -.
FT NON_TER 1
FT NON_TER 437 437
SQ SEQUENCE 437 AA; 50232 MW; BA3061373049C9CE CRC64;

Query Match 100.0%; Score 105; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQYEQHLYDYKKEG 19
|||||
Db 202 DQPKQYEQHLYDYKKEG 220

15

RESULT 13
Q9BHR0 ID Q9BHR0 PRELIMINARY; PRT; 437 AA.

AC Q9BHR0; 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1.

OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=NIGERIAN 035;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical
RT Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ408315; CAC34756.1; -.
FT NON_TER 1
FT NON_TER 437 437
SQ SEQUENCE 437 AA; 50039 MW; 625A5E6F26E01064 CRC64;

Query Match 100.0%; Score 105; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQYEQHLYDYKKEG 19
|||||
Db 202 DQPKQYEQHLYDYKKEG 220

RESULT 14

Q9BHQ9 ID Q9BHQ9 PRELIMINARY; PRT; 437 AA.

AC Q9BHQ9; 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA1.

OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=NIGERIAN 050;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical
RT Membrane Antigen 1 (AMA1).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ408316; CAC34757.1; -.
FT NON_TER 1
FT NON_TER 437 437
SQ SEQUENCE 437 AA; 50140 MW; 9FE785C026BEDC63 CRC64;

Query Match 100.0%; Score 105; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKQYEQHLYDYKKEG 19
|||||
Db 202 DQPKQYEQHLYDYKKEG 220

RESULT 15

Q9BHQ8 ID Q9BHQ8 PRELIMINARY; PRT; 437 AA.

```
AC Q9BHQ8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMAL.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NIGERIAN 050;
RA Polley S.D., Conway D.J.;
RT "Diversifying selection in domains of Plasmodium falciparum Apical
   Membrane Antigen 1 (AMAL).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ408317; CAC34758.1; -.
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 50144 MW; 991DB652BD8134F CRC64;

Query Match      100.0%; Score 105; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. NO. 3.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DQPKYEQHLDYKIREG 19
   |||
Db 202 DQPKYEQHLDYKIREG 220
   |||
```

Search completed: January 29, 2002, 11:12:14
Job time: 770 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:21:50 ; Search time 310.82 Seconds
(without alignments)
5.243 Million cell updates/sec

Title: US-09-763-397a-19

Perfect score: 131

Sequence: 1 EFTYMINFRGQNYWEHPYQKS 22

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
- 22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	131	100.0	22	21	AAV70295
2	131	100.0	350	21	AAV70278
3	131	100.0	622	10	AAV91632
4	126	96.2	622	13	AAV7532
5	126	96.2	622	16	AAV68840
6	52	39.7	506	21	AAV20572
7	49	37.4	512	21	AAV20573
8	46.5	35.5	346	13	AAV27576
9	46	35.1	133	21	AAV08691
10	46	35.1	133	21	AAV42699
11	46	35.1	151	20	AAV94252

12	46	35.1	230	21	AAV08690
13	46	35.1	230	21	AAV42698
14	46	35.1	255	22	AAV92532
15	46	35.1	262	20	AAV76587
16	45	34.4	124	22	AAU14709
17	45	34.4	168	22	AAV74619
18	45	34.4	168	22	AAV86883
19	45	34.4	171	18	AAW24980
20	45	34.4	474	12	AAV14676
21	45	34.4	631	10	AAV91139
22	45	34.4	631	10	AAV96143
23	45	34.4	631	11	AAV07350
24	45	34.4	631	11	AAV07969
25	45	34.4	631	19	AAW41226
26	45	34.4	644	22	AAV20490
27	45	34.4	660	11	AAV06420
28	45	34.4	660	22	AAV84607
29	45	34.4	663	19	AAW41111
30	45	34.4	663	19	AAW41227
31	44	33.6	10	11	AAV07369
32	44	33.6	140	19	AAW84149
33	44	33.6	140	20	AAV95518
34	44	33.6	140	20	AAV85128
35	44	33.6	140	21	AAV84696
36	44	33.6	140	21	AAV92611
37	44	33.6	376	22	AAV93070
38	44	33.6	509	18	AAV27249
39	44	33.6	518	22	AAV95260
40	44	33.6	933	22	AAV84151
41	43.5	33.2	122	21	AAV2675
42	43.5	33.2	379	22	AAV52464
43	43	32.8	181	22	AAV64952
44	43	32.8	219	21	AAV01400
45	43	32.8	302	21	AAV43408

ALIGNMENTS

RESULT 1	
AAV70295	
ID	AAV70295 standard; peptide: 22 AA.
XX	
AC	AAV70295;
XX	
DT	06-JUN-2000 (first entry)
XX	
DE	Plasmodium falciparum AMA-1 antigenic epitope, P603.
XX	
KW	Recombinant protein; CDC/NIAID/VAC-1; multivalent; malaria; vaccine;
KW	T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
KW	Circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
KW	Liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
KW	Apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
KW	EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
KW	Pf27; antiparasitic; prevention; anti-CDC/NIAID/VAC-1 antibody.
XX	
OS	Plasmodium falciparum.
XX	
PN	WO200011179-A1.
XX	
PD	02-MAR-2000.
XX	
PF	19-AUG-1999; 99WO-US18869.
XX	
PR	21-AUG-1998; 98US-0097703.
XX	
PA	(NAIM-) NAT INST IMMUNOLOGY.
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX	
PI	Lal AA, Shi YP, Hasnain SE;
XX	
DR	WPI; 2000-237654/20.

XX Novel recombinant protein as vaccine for treating malarial infection
 PT comprises antigenic peptides obtained from different stages of
 PT Plasmodium falciparum life cycle -
 XX
 XX Claim 2; Page 17; 52pp; English.
 XX
 XX The present sequence is the antigenic epitope P603, derived from
 CC apical membrane antigen-1 (AMA-1) of the asexual blood stage of
 CC Plasmodium falciparum. It is used in the construction of recombinant
 CC protein CDC/NiMALVAC-1, which is a multivalent, multistage malarial
 CC vaccine. The recombinant protein comprises, melittin signal peptide
 CC (His)6 tag, T-cell epitope from tetanus toxoid and 21 antigenic epitopes
 CC from circumsporozoite protein (CSP), sporozoite surface protein-2
 CC (SSP-2), liver stage antigen-1 (LSA-1), merozoite surface protein-1
 CC (MSP-1), MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding
 CC antigen-175 (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete
 CC specific antigen, Pf27. These epitopes were obtained at different stages
 CC of the life cycle of P. falciparum. CDC/NiMALVAC-1 vaccine has
 CC antiparasitic activity and can be used for treatment and prevention of
 CC malarial infections. Anti-CDC/NiMALVAC-1 antibodies can be used for
 CC detecting P. falciparum in biological samples.
 XX
 XX Sequence 22 AA;
 SQ
 Query Match 100.0%; Score 131; DB 21; Length 22;
 Best Local Similarity 100.0%; Pred. No. 7.9e-14;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EFTYMINFGRGQNYWEHPYQKS 22
 Db 1 eftyminfgrgnywehpyqks 22
 RESULT 2
 AAY70278
 ID AAY70278 standard; Protein; 350 AA.
 XX
 XX AAY70278;
 DT 06-JUN-2000 (first entry)
 DE Recombinant vaccine CDC/NiMALVAC-1.
 XX
 KW Recombinant protein; CDC/NiMALVAC-1; multivalent; malaria; vaccine;
 KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
 KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
 KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
 KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
 KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
 KW Pf27; antiparasitic; prevention; anti-CDC/NiMALVAC-1 antibody;
 KW honey bee.
 XX
 OS Chimeric - Apis sp.
 OS Chimeric - Clostridium tetani.
 OS Chimeric - Plasmodium falciparum.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..22
 FT /label= Melittin_signal_peptide
 FT /note= "Derived from Honey Bee"
 FT Protein 23..350
 FT /label= Mature_CDC/NiMALVAC-1
 FT /note= "Recombinant multivalent malarial vaccine"
 XX
 XX WO200011179-A1.
 PN
 XX 02-MAR-2000.
 PD
 XX 19-AUG-1999; 99WO-US18869.
 PF
 XX 21-AUG-1998; 98US-0097703.

XX (NAlM-) NAT INST IMMUNOLOGY.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Lal AA, Shi YP, Hasnain SE;
 XX
 XX WPI; 2000-237654/20.
 DR N-PSDB; AA251336.
 XX
 XX Novel recombinant protein as vaccine for treating malarial infection
 PT comprises antigenic peptides obtained from different stages of
 PT Plasmodium falciparum life cycle -
 XX
 XX Claim 3; Page 43-44; 52pp; English.
 XX
 XX The present sequence is that of recombinant protein CDC/NiMALVAC-1,
 CC which is a multivalent, multistage malarial vaccine. The recombinant
 CC protein comprises, melittin signal peptide, (His)6 tag, T-cell epitope
 CC from tetanus toxoid and 21 antigenic epitopes from circumsporozoite
 CC protein (CSP), sporozoite surface protein-2 (SSP-2), liver stage
 CC antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1), MSP-2, apical
 CC membrane antigen-1 (AMA-1), erythrocyte binding antigen-175 (EBA-175),
 CC rhoptry associated protein-1 (RAP-1) and gamete specific antigen, Pf27.
 CC These epitopes were obtained at different stages of the life cycle of
 CC Plasmodium falciparum. CDC/NiMALVAC-1 vaccine has antiparasitic
 CC activity and can be used for treatment and prevention of malarial
 CC infections. Anti-CDC/NiMALVAC-1 antibodies can be used for detecting
 CC P. falciparum in biological samples.
 XX
 XX Sequence 350 AA;
 SQ
 Query Match 100.0%; Score 131; DB 21; Length 350;
 Best Local Similarity 100.0%; Pred. No. 1.9e-12;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EFTYMINFGRGQNYWEHPYQKS 22
 Db 242 eftyminfgrgnywehpyqks 263
 RESULT 3
 AAP91632
 ID AAP91632 standard; protein; 622 AA.
 XX
 XX AAP91632;
 DT 25-JAN-1990 (first entry)
 XX
 DE Rhoptry membrane antigen-1.
 XX
 KW Malaria; rhoptry membrane antigen-1; antibodies.
 XX
 OS Plasmodium falciparum D10.
 XX
 XX WO8907645-A.
 PN
 XX 24-AUG-1989.
 PD
 XX 10-FEB-1989; 89WO-AU00056.
 PF
 XX 12-FEB-1988; 88AU-0006743.
 PR
 XX (SARA) SARAMANE PTY LTD.
 PA
 XX Peterson MG, Crewther PE, Smythe JA, Marshall VM, Silva A;
 PI
 XX WPI; 1989-263714/36.
 DR N-PSDB; AAN90703.
 XX
 XX Rhoptry membrane antigen of Plasmodium falciparum
 PT - used for producing antibodies and in immunisation,
 PT diagnostic and treatment methods for malaria.
 PT


```

XX PS Claim 1; Fig 3; 46pp; English.
XX
CC RMA-1 can generate an immune response to malaria, and antibodies which
CC can inhibit growth of the parasite. RMA-1 initially has mol. wt. 80 kD.
XX
SQ Sequence 622 AA;

Query Match 100.0%; Score 131; DB 10; Length 622;
Best Local Similarity 100.0%; Pred. No. 3.6e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EFTYMINFGNGQNYWEHPYQKS 22
Db 14 eftyminfgrgnywehpyqks 35

RESULT 4
AAR27532
ID AAR27532 standard; Protein; 622 AA.
XX
AC AAR27532;
XX
DT 08-MAR-1993 (first entry)
XX
DE Plasmodium falciparum AMA1 antigen.
XX
KW Recombinant poxvirus; antimalarial vaccine; malaria; immunise;
KW immunogen.
XX
OS Plasmodium falciparum.
XX
PN W09216616-A.
XX
PD 01-OCT-1992.
XX
PF 19-MAR-1992; 92WO-US02207.
XX
PR 20-MAR-1991; 91US-0672183.
PR 18-MAR-1992; 92US-0852305.
XX
PA (VIRO-) VIROGENETICS CORP.
XX
PI De Taisne C, Paoletti E, Tine JA;
XX
DR WPI; 1992-349203/42.
DR N-PSDB; AAQ29189.
XX
PT Recombinant poxvirus - contg. Plasmodium DNA, useful as
PT antimalarial vaccine
XX
PS Example 4; Fig 5; 74pp; English.
XX
CC This sequence is the Plasmodium falciparum AMA1 antigen.
CC cDNA encoding it was cloned into vaccinia donor plasmids
CC before being inserted into the vaccinia virus to be used in a
CC vaccine to stimulate an antimalarial immunological response, or for
CC in vitro prodn. of gene prods. for use as immunogens. As plasmodium
CC genes are conserved among P. falciparum strains, they are widely
CC effective in a vaccine.
XX
SQ Sequence 622 AA;

Query Match 96.2%; Score 126; DB 13; Length 622;
Best Local Similarity 95.5%; Pred. No. 2.2e-11;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EFTYMINFGNGQNYWEHPYQKS 22
Db 14 eftyminfgrgnywehpyqks 35

XX PS Claim 1; Fig 3; 46pp; English.
XX
CC RMA-1 can generate an immune response to malaria, and antibodies which
CC can inhibit growth of the parasite. RMA-1 initially has mol. wt. 80 kD.
XX
SQ Sequence 622 AA;

Query Match 100.0%; Score 131; DB 10; Length 622;
Best Local Similarity 100.0%; Pred. No. 3.6e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EFTYMINFGNGQNYWEHPYQKS 22
Db 14 eftyminfgrgnywehpyqks 35

RESULT 4
AAR27532
ID AAR27532 standard; Protein; 622 AA.
XX
AC AAR27532;
XX
DT 08-MAR-1993 (first entry)
XX
DE Plasmodium falciparum AMA1 antigen.
XX
KW Recombinant poxvirus; antimalarial vaccine; malaria; immunise;
KW immunogen.
XX
OS Plasmodium falciparum.
XX
PN W09216616-A.
XX
PD 01-OCT-1992.
XX
PF 19-MAR-1992; 92WO-US02207.
XX
PR 20-MAR-1991; 91US-0672183.
PR 18-MAR-1992; 92US-0852305.
XX
PA (VIRO-) VIROGENETICS CORP.
XX
PI De Taisne C, Paoletti E, Tine JA;
XX
DR WPI; 1992-349203/42.
DR N-PSDB; AAQ29189.
XX
PT Recombinant poxvirus - contg. Plasmodium DNA, useful as
PT antimalarial vaccine
XX
PS Example 4; Fig 5; 74pp; English.
XX
CC This sequence is the Plasmodium falciparum AMA1 antigen.
CC cDNA encoding it was cloned into vaccinia donor plasmids
CC before being inserted into the vaccinia virus to be used in a
CC vaccine to stimulate an antimalarial immunological response, or for
CC in vitro prodn. of gene prods. for use as immunogens. As plasmodium
CC genes are conserved among P. falciparum strains, they are widely
CC effective in a vaccine.
XX
SQ Sequence 622 AA;

Query Match 96.2%; Score 126; DB 13; Length 622;
Best Local Similarity 95.5%; Pred. No. 2.2e-11;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EFTYMINFGNGQNYWEHPYQKS 22
Db 14 eftyminfgrgnywehpyqks 35

```

```

RESULT 5
AAR68840
ID AAR68840 standard; Protein; 622 AA.
XX
AC AAR68840;
XX
DT 24-AUG-1995 (first entry)
XX
DE Plasmodium falciparum AMA-1 gene protein.
XX
KW Plasmodium falciparum AMA-1 gene; recombinant poxvirus;
KW multicomponent multistage malarial vaccines; Immunogens;
KW malaria diagnosis.
XX
OS Plasmodium falciparum (3D7).
XX
PN W09428930-A.
XX
PD 22-DEC-1994.
XX
PF 10-JUN-1994; 94WO-US06652.
XX
PR 11-JUN-1993; 93US-0075783.
PR 09-JUN-1994; 94US-0257073.
XX
PA (VIRO-) VIROGENETICS CORP.
XX
PI De Taisne C, Paoletti E, Tine JA;
XX
DR WPI; 1995-036113/05.
DR N-PSDB; AAQ80910.
XX
PT Recombinant poxvirus contg. Plasmodium DNA in non-essential
PT region - useful in vaccines against malaria and for prodn. of
PT Plasmodium immunogens
XX
PS Claim 3; Fig 5; 183pp; English.
XX
CC AAQ80910 encodes AAR68840 the P. falciparum AMA-1 gene product. New
CC recombinant poxviruses containing either the SERA, ABRA, PfhsP70,
CC AMA-1, PfS25, PfS16, CSP, PfSSP2, LSA-1, LSA-1 repeats, MSA-1,
CC MSA-1 (N-terminal p83 or C-terminal gp42) genes, or a combination
CC of these in non-essential regions of their genomes are claimed.
CC These poxviruses (pref. with a virulence reducing genomic
CC deletion or disruption) can be used as vaccines against malaria
CC and for the prodn. of Plasmodium immunogens. These viruses
CC provide multicomponent, multistage vaccines due to their expression
CC of sporozite, liver stage, blood stage and sexual stage proteins.
XX
SQ Sequence 622 AA;

Query Match 96.2%; Score 126; DB 16; Length 622;
Best Local Similarity 95.5%; Pred. No. 2.2e-11;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EFTYMINFGNGQNYWEHPYQKS 22
Db 14 eftyminfgrgnywehpyqks 35

RESULT 6
AAB20572
ID AAB20572 standard; Protein; 506 AA.
XX
AC AAB20572;
XX
DT 08-DEC-2000 (first entry)
XX
DE Antirrhinum majus flavone synthase SEQ ID NO:2.
XX
KW Flavone; flavanone; flavone synthase; plant; flower colour;

```

KW bacterial resistance; nitrogen fixing bacteria; protection;
KW ultraviolet radiation.

XX Antirrhinum majus.

XX WO2000044907-A1.

XX PD 03-AUG-2000.

XX PF 28-JAN-2000; 2000WO-JP00490.

XX PR 29-JAN-1999; 99JP-0022427.

XX PR 19-JUL-1999; 99JP-0205229.

XX PA (SUNR) SUNTORY LTD.

XX PI Mizutani M, Tanaka Y, Kusumi T, Ayabe S, Akashi T;

XX DR WPI: 2000-543394/49.

XX DR N-PSDB; AAA87996.

XX Gene encoding a flavone synthase for production of Antirrhinum and
XX other plants with modified flower colour

XX PS Claim 3; Page 31-34; 54pp; Japanese.

XX CC The present sequence represents a protein which has flavone synthase
XX activity for converting flavanone directly into flavone. Flavone
XX synthase can be used in the production of plant varieties with modified
XX flower colour, improved resistance to bacteria, improved interaction
XX with nitrogen fixing bacteria, or increased protection against
XX ultraviolet radiation.

XX SQ Sequence 506 AA;

Query Match 39.7%; Score 52; DB 21; Length 506;

Best Local Similarity 42.1%; Pred. No. 6.4;

Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 FTYMINFGRCQNYWEHPYQ 20

I : : || ||||| :

Db 391 fvnlsgrnpywspme 409

RESULT 7

AAB20573

ID AAB20573 standard; Protein; 512 AA.

XX AC AAB20573;

XX DT 08-DEC-2000 (first entry)

XX DE Torenia hybrida flavone synthase SEQ ID NO:4.

XX KW Flavone; flavanone; flavone synthase; plant; flower colour;
XX KW bacterial resistance; nitrogen fixing bacteria; protection;
XX KW ultraviolet radiation.

XX OS Torenia hybrida.

XX PN WO2000044907-A1.

XX PD 03-AUG-2000.

XX PF 28-JAN-2000; 2000WO-JP00490.

XX PR 29-JAN-1999; 99JP-0022427.

XX PR 19-JUL-1999; 99JP-0205229.

XX PA (SUNR) SUNTORY LTD.

XX PI Mizutani M, Tanaka Y, Kusumi T, Ayabe S, Akashi T;

XX DR WPI: 2000-543394/49.

XX DR N-PSDB; AAA87997.

XX Gene encoding a flavone synthase for production of Antirrhinum and
XX other plants with modified flower colour

XX PS Claim 3; Page 38-40; 54pp; Japanese.

XX CC The present sequence represents a protein which has flavone synthase
XX activity for converting flavanone directly into flavone. Flavone
XX synthase can be used in the production of plant varieties with modified
XX flower colour, improved resistance to bacteria, improved interaction
XX with nitrogen fixing bacteria, or increased protection against
XX ultraviolet radiation.

XX SQ Sequence 512 AA;

Query Match 37.4%; Score 49; DB 21; Length 512;

Best Local Similarity 42.1%; Pred. No. 19;

Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 FTYMINFGRCQNYWEHPYQ 20

I : : || ||||| :

Db 396 fvnlsgrnpywspme 414

RESULT 8

AAR27576

ID AAR27576 standard; Protein; 346 AA.

XX AC AAR27576;

XX DT 20-MAY-1998 (first entry)

XX DE ABN-A from A. niger.

XX KW Endo 1,5-alpha-L-arabinase A; arabinan; juice; fruit; vegetable; haze;
XX KW yield; wood pulp; lignins; terpenoids; paper processing.

XX OS Aspergillus niger.

XX FH Key Location/Qualifiers

FT peptide 1..19

FT /note= "signal peptide"

FT protein 20..346

FT /note= "mature protein"

XX PN EP506190-A.

XX PD 30-SEP-1992.

XX PF 20-MAR-1992; 92EP-0200818.

XX PR 27-MAR-1991; 91EP-0200720.

XX PA (KONN) GIST-BROCADES NV.

XX PI Andreoli PM, Bakhuis JG, Coutel Y, De Graaff LH, Flippin MJA;

XX PI Hardera, Van der Veen P, Van Heuvel M, Visser J;

XX DR WPI: 1992-325492/40.

XX DR P-PDSB; Q29231.

XX PT Arabinan-degrading enzymes and genes - isolated from fungi and
XX PT having lower pH optima than arabinan-degrading enzymes of
XX bacterial origin

XX PS Example 16; Fig 20; 75pp; English.

XX CC The protein sequence was deduced from the DNA sequence obt'd. by
XX screening an A. niger N400 genomic library with a probe which

CC obtd. from a primary screen of *A. niger* DNA with anti-ABN A antibodies.
CC ABN-A has arabinan-degrading activity and has the ability to cleave
CC (1-2), (1-3) or (1-5) alpha L-arabinosidic linkages and can cleave the
CC 1-6 linkage between the terminal arabinofuranosyl unit and the
CC intermediate glucosyl unit of monoterpenyl alpha-L-arabinofuranosyl
CC glucosides. The enzyme has a lower pH optimum than those of bacterial
CC origin. The enzyme may be used in the prodn. of concd. juices, e.g.
CC fruit or vegetable, to eliminate arabinan haze, to increase the yield
CC of juice from pulp, to improve the filterability and reduce the
CC viscosity of juices, to assist in the liberation of aroma cpds. from
CC substrates such as juices or wines, to improve the in vivo uptake and
CC utilisation of plant nutrients by animals, to provide more efficient
CC removal of both lignins and terpenoids in wood pulp and paper processing
CC and in the hydrolysis of sugar beet pulp, sycamore or gum arabic or
CC agricultural residues such as wheat-straw.
CC See also AAR27571-75.
XX
SQ Sequence 346 AA;

Query Match 35.5%; Score 46.5; DB 13; Length 346;
Best Local Similarity 50.08; Pred. No. 30;
Matches 9; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

Qy 3 TYMINFGRGQNYWEHPYQ 20
||| ||| :|: ||
Db 158 tyinfg---sfwdidiy 172

RESULT 9

AAG08691
ID AAG08691 standard; Protein: 133 AA.

XX AC AAG08691;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 6329.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 28-APR-1999; 99US-0130891.

XX PR 30-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 04-MAY-1999; 99US-0132407.

XX PR 05-MAY-1999; 99US-0132484.

XX PR 06-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

XX PR 07-MAY-1999; 99US-0132487.

XX PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.

```
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 35.1%; Score 46; DB 21; Length 133;
Best Local Similarity 46.2%; Pred. No. 12;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 GRGONYWEHPYOK 21
| |: :||| :||
Db 13 gsgeafwehewek 25

RESULT 10
AAG42699
ID AAG42699 standard; Protein; 133 AA.
XX
AC AAG42699;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 53280.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
```

PR 20-MAY-1999;	99US-0135124.	PR 02-AUG-1999;	99US-0146386.
PR 21-MAY-1999;	99US-0135353.	PR 02-AUG-1999;	99US-0146388.
PR 24-MAY-1999;	99US-0135629.	PR 02-AUG-1999;	99US-0146389.
PR 25-MAY-1999;	99US-0136021.	PR 03-AUG-1999;	99US-0147038.
PR 27-MAY-1999;	99US-0136392.	PR 04-AUG-1999;	99US-0147204.
PR 28-MAY-1999;	99US-0136782.	PR 04-AUG-1999;	99US-0147302.
PR 01-JUN-1999;	99US-0137222.	PR 05-AUG-1999;	99US-0147192.
PR 03-JUN-1999;	99US-0137528.	PR 05-AUG-1999;	99US-0147260.
PR 04-JUN-1999;	99US-0137502.	PR 06-AUG-1999;	99US-0147303.
PR 07-JUN-1999;	99US-0137724.	PR 06-AUG-1999;	99US-0147416.
PR 08-JUN-1999;	99US-0138094.	PR 09-AUG-1999;	99US-0147493.
PR 10-JUN-1999;	99US-0138540.	PR 09-AUG-1999;	99US-0147935.
PR 10-JUN-1999;	99US-0138847.	PR 10-AUG-1999;	99US-0148171.
PR 14-JUN-1999;	99US-0139119.	PR 11-AUG-1999;	99US-0148319.
PR 16-JUN-1999;	99US-0139452.	PR 12-AUG-1999;	99US-0148341.
PR 17-JUN-1999;	99US-0139453.	PR 13-AUG-1999;	99US-0148565.
PR 17-JUN-1999;	99US-0139492.	PR 13-AUG-1999;	99US-0148684.
PR 18-JUN-1999;	99US-0139454.	PR 16-AUG-1999;	99US-0149368.
PR 18-JUN-1999;	99US-0139455.	PR 17-AUG-1999;	99US-0149175.
PR 18-JUN-1999;	99US-0139456.	PR 18-AUG-1999;	99US-0149426.
PR 18-JUN-1999;	99US-0139457.	PR 20-AUG-1999;	99US-0149722.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139459.	PR 20-AUG-1999;	99US-0149929.
PR 18-JUN-1999;	99US-0139460.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139461.	PR 23-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139462.	PR 25-AUG-1999;	99US-0150566.
PR 18-JUN-1999;	99US-0139463.	PR 26-AUG-1999;	99US-0150884.
PR 18-JUN-1999;	99US-0139750.	PR 27-AUG-1999;	99US-0151065.
PR 18-JUN-1999;	99US-0139763.	PR 27-AUG-1999;	99US-0151066.
PR 21-JUN-1999;	99US-0139817.	PR 27-AUG-1999;	99US-0151080.
PR 22-JUN-1999;	99US-0139899.	PR 30-AUG-1999;	99US-0151303.
PR 23-JUN-1999;	99US-0140353.	PR 31-AUG-1999;	99US-0151438.
PR 23-JUN-1999;	99US-0140354.	PR 01-SEP-1999;	99US-0151930.
PR 24-JUN-1999;	99US-0140699.	PR 07-SEP-1999;	99US-0152363.
PR 28-JUN-1999;	99US-0140823.	PR 10-SEP-1999;	99US-0153070.
PR 29-JUN-1999;	99US-0140991.	PR 13-SEP-1999;	99US-0153758.
PR 30-JUN-1999;	99US-0141287.	PR 15-SEP-1999;	99US-0154018.
PR 01-JUL-1999;	99US-0141842.	PR 16-SEP-1999;	99US-0154039.
PR 01-JUL-1999;	99US-0142154.	PR 20-SEP-1999;	99US-0154779.
PR 02-JUL-1999;	99US-0142055.	PR 22-SEP-1999;	99US-0155139.
PR 06-JUL-1999;	99US-0142390.	PR 23-SEP-1999;	99US-0155486.
PR 08-JUL-1999;	99US-0142803.	PR 24-SEP-1999;	99US-0155659.
PR 09-JUL-1999;	99US-0142920.	PR 28-SEP-1999;	99US-0156458.
PR 12-JUL-1999;	99US-0142977.	PR 29-SEP-1999;	99US-0156596.
PR 13-JUL-1999;	99US-0143542.	PR 04-OCT-1999;	99US-0157117.
PR 14-JUL-1999;	99US-0143624.	PR 05-OCT-1999;	99US-0157753.
PR 15-JUL-1999;	99US-0144005.	PR 06-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144085.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144086.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144325.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144884.	PR 14-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144814.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145276.	PR 25-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145918.	PR 26-OCT-1999;	99US-0161359.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161360.
PR 28-JUL-1999;	99US-0145951.		

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 35.1%; Score 46; DB 21; Length 133;
Best Local Similarity 46.2%; Pred. No. 12;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 GRGQNYWEHPYQK 21
| | : | | | : | |
Db 13 gsgeafwehewek 25

RESULT 11
AAW94252
ID AAW94252 standard; Protein; 151 AA.
XX
AC AAW94252;
XX
DT 16-APR-1999 (first entry)
XX
DE Human formin binding protein (FBPhu).
XX
KW Formin binding protein; FBPhu; human; developmental disorder; anaemia;
KW renal tubular acidosis; Cushing's syndrome; epilepsy; spina bifida;
KW glaucoma; cancer; tumour; leukaemia; lymphoma; immune suppression;
KW autoimmunity; inflammation; acquired immune deficiency syndrome; AIDS;
KW allergy; asthma; ulcerative colitis; multiple sclerosis; osteoporosis;
KW microbial; viral; infection; chromosomal mapping.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 151
FT /label= "unknown"
FT /note= "encoded by NGA"
XX
XX W09856912-A1.
XX
PD 17-DEC-1998.
XX
XX
PF 08-JUN-1998; 98WO-US11939.
XX
XX 11-JUN-1997; 97US-0872783.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Hillman JL, Lal P;
XX
XX WPI; 1999-070323/06.
DR N-PSDB; AAX05111.
XX
XX New human formin-binding protein - useful in the treatment and
PT diagnosis of cancer, immune system diseases and developmental
PT disorders
XX
XX Claim 1; Fig 1A-B; 61pp; English.
XX
XX This represents a human formin binding protein (FBPhu). Host cells
CC containing a vector comprising the FBPhu nucleic acid are used for the
CC recombinant production of the protein. Recombinant FBPhu polypeptides
CC are used for generation of Ab and for screening for specific modulators.
CC The FBPhu polypeptide and its agonists (optionally expressed from gene
CC therapy vectors) are used to treat developmental disorders, e.g renal
CC tubular acidosis, anaemia, Cushing's syndrome, epilepsy, spina bifida,
CC congenital glaucoma and many others. Antagonists are used to treat a
CC wide range of cancers (solid tumours, leukaemia and lymphoma), also
CC diseases that involve immune suppression, autoimmunity or inflammation
CC (typical of many examples are acquired immune deficiency syndrome,
CC allergy, asthma, ulcerative colitis, multiple sclerosis, osteoporosis,
CC microbial or viral infection). Fragments of FBPhu nucleic acid are used,

CC for diagnosis and monitoring, to detect FBPhu-encoding nucleic acid (in
CC standard hybridisation and/or amplification assays); as therapeutic
CC antisense, triplex-forming or ribozyme molecules; to identify genetic
CC variation, mutations etc., and for chromosomal mapping.

XX
SQ Sequence 151 AA;

Query Match 35.1%; Score 46; DB 20; Length 151;
Best Local Similarity 47.1%; Pred. No. 14;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 FTYMINFGRCQNYWEHP 18
| | | | | | : | | | |
Db 111 ftyyntetgesrwekp 127

RESULT 12
AAG08690
ID AAG08690 standard; Protein; 230 AA.
XX
AC AAG08690;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 6328.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.

PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0158293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161992.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147204.	PR	29-OCT-1999;	99US-0162142.

Query Match 35.1%; Score 46; DB 21; Length 230;

Best Local Similarity 46.2%; Pred. No. 22;

Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 GRGONYWEHPYOK 21

Db 110 gsgaeafwehewek 122

RESULT 13

AAG42698
ID AAG42698 standard; Protein; 230 AA.

XX AAG42698;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 53279.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 18-MAY-1999; 99US-0134370.

PR 19-MAY-1999; 99US-0134768.

PR 20-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.

Job time: 426 sec

QY 2 FTYMINEGRGQNYWEHP 18
| | | | | : | | |
Db 178 ftyyintetgesrwekp 194

RESULT 15

AA776587
ID AAY76587 standard; Protein; 262 AA.

XX AC AAY76587;

XX DT 10-APR-2000 (first entry)

XX DE Human ovarian tumor EST fragment encoded protein 83.

XX KW Expressed sequence tag; EST; human; ovarian tumor; anticancer;
gene therapy; treatment.

XX OS Homo sapiens.

XX PN DE19817557-A1.

XX PD 21-OCT-1999.

XX PF 09-APR-1998; 98DE-1017557.

XX PR 09-APR-1998; 98DE-1017557.

XX PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;

XX DR WPI; 1999-591920/51.

XX DR N-PSDB; AAZ77485.

XX PT New nucleic acid sequences expressed in ovarian, and some other, cancer
tissues, and derived polypeptides, for treatment of ovarian cancer and
identification of therapeutic agents

XX PS Claim 25; Page 277; 310pp; German.

XX CC This invention describes novel nucleic acid (CDNA) sequences (A) which
have anticancer activity and are highly expressed in ovarian tumor
tissue (and some also in testis and breast cancer tissue). The products
of the invention can be used for gene therapy. (A) are used (i) for
recombinant expression of polypeptides (B) and (ii) to isolate complete
genes. (B) are used (i) to identify agents suitable for treatment of
ovarian cancer; (ii) directly for treating this form of cancer
(including expression from gene therapy vectors) and (iii) for generation
of specific antibodies. (A) are identified by assembling ESTs (expressed
sequence tags) from a particular tissue type before comparison of
expression patterns. This allows a significantly longer fragment of the
gene to be revealed, so should reduce the number of failures associated
with the fact that ESTs from different libraries may represent different
parts of the same unknown gene, distorting the estimated frequency of
occurrence in a particular tissue. AAY76505-Y76638 represent protein
fragments encoded by the human ovarian tumor cDNA library derived EST
fragments represented in AAZ77450-77752.

XX SQ Sequence 262 AA;

Query Match 35.1%; Score 46; DB 20; Length 262;
Best Local Similarity 47.1%; Pred. No. 26;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 FTYMINEGRGQNYWEHP 18
| | | | | : | | |
Db 79 ftyyintetgesrwekp 95

Search completed: January 29, 2002, 10:21:50

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:24:10 ; Search time 133.18 seconds
(without alignments)
3.717 Million cell updates/sec

Title: US-09-763-397A-19
Perfect score: 131
Sequence: 1 EFTYMINFGKQNYWHPYQKS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles!.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	126	96.2	628	1	US-08-257-073-9
2	46.5	35.5	346	2	US-07-952-853-24
3	46.5	35.5	346	2	US-08-914-848-24
4	46	35.1	151	2	US-08-872-783-1
5	45	34.4	168	4	US-08-444-628-9
6	45	34.4	168	4	US-08-357-820-9
7	45	34.4	171	1	US-08-303-270-1
8	45	34.4	631	4	US-08-448-489-17
9	45	34.4	660	3	US-08-704-711A-18
10	43	32.8	67	2	US-08-872-783-3
11	43	32.8	114	2	US-08-741-437-3
12	43	32.8	114	2	US-09-134-593-3
13	43	32.8	219	2	US-08-773-368-3
14	43	32.8	219	3	US-09-199-887-3
15	43	32.8	266	2	US-08-773-368-4
16	43	32.8	266	3	US-09-199-887-4
17	42	32.1	171	1	US-08-313-075A-32
18	42	32.1	542	1	US-07-814-964-13
19	42	32.1	542	1	US-08-258-442-13
20	42	32.1	542	1	US-08-328-809-8
21	42	32.1	542	5	PCT-US92-11107-13
22	42	32.1	610	2	US-08-724-394A-5
23	41.5	31.7	324	2	US-08-816-755-2
24	41.5	31.7	324	4	US-09-090-673-2
25	41.5	31.7	532	4	US-09-294-841-2
26	41.5	31.7	579	3	US-08-704-711A-1
27	41.5	31.7	582	3	US-08-704-711A-2

RESULT 1
US-08-257-073-9
; Sequence 9, Application US/08257073
; Patent No. 5766597
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: de Taisne, Charles
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue, 25th Floor
; CITY: New York
; STATE: New York
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,073
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/075,783
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/852,305
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,183
; FILING DATE: 20-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2570
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 628 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

ALIGNMENTS

28	41.5	31.7	582	4	US-08-448-489-1	Sequence 1, Appli
29	41.5	31.7	582	4	US-09-211-704A-9	Sequence 9, Appli
30	41	31.3	119	2	US-08-561-521-13	Sequence 13, Appl
31	41	31.3	119	5	PCT-US95-01219-13	Sequence 13, Appl
32	41	31.3	1835	3	US-08-836-325-15	Sequence 13, Appl
33	41	31.3	1969	3	US-08-836-325-16	Sequence 16, Appl
34	41	31.3	1984	3	US-08-836-325-10	Sequence 10, Appl
35	41	31.3	1989	3	US-08-836-325-11	Sequence 11, Appl
36	41	31.3	1989	3	US-08-836-325-12	Sequence 12, Appl
37	40.5	30.9	42	3	US-08-924-330A-8	Sequence 8, Appli
38	40.5	30.9	42	4	US-09-138-721-8	Sequence 8, Appli
39	40.5	30.9	123	2	US-08-937-972-1	Sequence 1, Appli
40	40.5	30.9	488	1	US-07-794-393-2	Sequence 2, Appli
41	40.5	30.9	488	1	US-08-001-711-2	Sequence 2, Appli
42	40.5	30.9	488	3	US-08-704-711A-22	Sequence 22, Appl
43	40.5	30.9	489	4	US-08-448-489-11	Sequence 11, Appl
44	40	30.5	115	2	US-08-308-494A-17	Sequence 17, Appl
45	40	30.5	229	1	US-08-158-682A-2	Sequence 2, Appli

FRAGMENT TYPE: internal
US-08-257-073-9

Query Match 96.2%; Score 126; DB 1; Length 628;
Best Local Similarity 95.5%; Pred. No. 4.6e-12;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EFTYMINFGKQNYWEHPYQKS 22
|||||
Db 14 EFTYMINFGKQNYWEHPYQNS 35

RESULT 2

US-07-952-853-24
; Sequence 24, Application US/07952853
; Patent No. 5863783

GENERAL INFORMATION:

APPLICANT: Van Heuvel, Margaretha
APPLICANT: Bakhuis, Janna G.
APPLICANT: Coutel, Yves
APPLICANT: Harder, Abraham
APPLICANT: De Graaff, Leendert H.
APPLICANT: Flippin, Michel J. A.
APPLICANT: Van Der Veen, Peter
APPLICANT: Visser, Jacob
APPLICANT: Andreoli, Peter M.
TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA
TITLE OF INVENTION: MOLECULES
TITLE OF INVENTION: ENCODING ARABINAN-DEGRADING ENZYMES OF FUNGAL
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/952,853
FILING DATE: 19921125
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 246152003500
TELEPHONE: 415-813-5600
TELEFAX: 415-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 346 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-952-853-24

Query Match 35.5%; Score 46.5; DB 2; Length 346;
Best Local Similarity 50.0%; Pred. No. 8.4;
Matches 9; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

QY 3 TYMINFGKQNYWEHPYQ 20
|||||
Db 158 TYINFG---SFWDIIYQ 172

RESULT 3

US-08-914-848-24

; Sequence 24, Application US/08914848
; Patent No. 5989887

GENERAL INFORMATION:

APPLICANT: Van Heuvel, Margaretha
APPLICANT: Bakhuis, Janna G.
APPLICANT: Coutel, Yves
APPLICANT: Harder, Abraham
APPLICANT: De Graaff, Leendert H.
APPLICANT: Flippin, Michel J. A.
APPLICANT: Van Der Veen, Peter
APPLICANT: Visser, Jacob
APPLICANT: Andreoli, Peter M.
TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA
TITLE OF INVENTION: MOLECULES ENCODING ARABINAN-DEGRADING ENZYMES OF FUNGAL
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/914,848
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/952,853
FILING DATE: 25-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 246152003500
TELEPHONE: 415-813-5600
TELEFAX: 415-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 346 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-914-848-24

Query Match 35.5%; Score 46.5; DB 2; Length 346;
Best Local Similarity 50.0%; Pred. No. 8.4;
Matches 9; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

QY 3 TYMINFGKQNYWEHPYQ 20
|||||
Db 158 TYINFG---SFWDIIYQ 172

RESULT 4

US-08-872-783-1

; Sequence 1, Application US/08872783
; Patent No. 5856717

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
TITLE OF INVENTION: HUMAN FORMIN BINDING PROTEIN
NUMBER OF SEQUENCES: 5

```

RESULT      5
US-08-444-628-9
; Sequence 9, Application US/08444628
; Patent No. 6184021
; GENERAL INFORMATION:
; APPLICANT: Senior, Robert M.
; TITLE OF INVENTION: Catalytically-active Gelatinase Mutant
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer, Monsanto/Searle, A3SG
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,628
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

RESULT 6
US-08-357-820-9
; Sequence 9, Application US/08357820
; Patent No. 6194189
; GENERAL INFORMATION:
; APPLICANT: Senior, Robert M.
; TITLE OF INVENTION: Catalytically-active Gelatinase Mutant
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:

```

/ FILING DATE:
/ CLASSIFICATION: 800
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Meyer, Scott J.
/ REGISTRATION NUMBER: 25,275
/ REFERENCE/DOCKET NUMBER: WU-20855
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (314)694-3117
/ TELEFAX: (314)694-5435
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 168 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-08-357-820-9

```

Query Match 34.4%; Score 45; DB 4; Length 168;
Best Local Similarity 58.8%; Pred. No. 6.5;
Matches 10; Conservative 0; Mismatches 3; Indels

QY 1 EFTYMINFGKQNYWEH 17
| | | | | | | | | |
Db 57 EADIMINFGKQNYWEH 69

RESULT 7

US-08-303-270-1
; Sequence 1, Application US/08303270
; Patent No. 5646027
; GENERAL INFORMATION:
; APPLICANT: Ye, Qi-Zhuang
; APPLICANT: Johnson, Linda L.
; APPLICANT: Hupe, Donald J.
; TITLE OF INVENTION: Process for the Production of
; TITLE OF INVENTION: Gelatinase Catalytic Domain Protein
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Rd.
; CITY: Ann Arbor
; STATE: MI
; COUNTRY: US
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/303,270
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Tinney, Francis J.
; REGISTRATION NUMBER: 33,069
; REFERENCE/DOCKET NUMBER: 5120-FJT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 313 996-7295
; TELEFAX: 313 996-1553
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-303-270-1

Query Match 34.4%; Score 45; DB 1; Length 171;
Best Local Similarity 58.8%; Pred. No. 6.6;
Matches 10; Conservative 0; Mismatches 3; Indels 4; Gaps 1;

QY 1 EFTYMINFGKQNYWEH 17
| | | | | | | | | |
Db 60 EADIMINFGKQNYWEH 72

RESULT 8

US-08-448-489-17
; Sequence 17, Application US/08448489
; Patent No. 6184022
; GENERAL INFORMATION:
; APPLICANT: SEIKI, Motoharu
; APPLICANT: SATO, Hiroshi
; APPLICANT: SHINAGAWA, Akira
; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
; FILE REFERENCE: 55-290P
; CURRENT APPLICATION NUMBER: US/08/448,489
; CURRENT FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17

Query Match 34.4%; Score 45; DB 3; Length 660;
Best Local Similarity 58.8%; Pred. No. 30;
Matches 10; Conservative 0; Mismatches 3; Indels 4; Gaps 1;

; LENGTH: 631
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Known Member of
; OTHER INFORMATION: Matrix Metalloproteinase Family
US-08-448-489-17

Query Match 34.4%; Score 45; DB 4; Length 631;
Best Local Similarity 58.8%; Pred. No. 28;
Matches 10; Conservative 0; Mismatches 3; Indels 4; Gaps 1;

QY 1 EFTYMINFGKQNYWEH 17

Db 137 EADIMINFGKQNYWEH 149

RESULT 9

US-08-704-711A-18
; Sequence 18, Application US/08704711A
; Patent No. 6114159
; GENERAL INFORMATION:
; APPLICANT: WILL, Horst
; APPLICANT: HINZMANN, Bernd
; TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
; TITLE OF INVENTION: METALLOPROTEASES, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/704,711A
; FILING DATE: 20-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/DE95/00357
; FILING DATE: 17-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4438838.1
; FILING DATE: 21-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4409663.1
; FILING DATE: 17-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: GRANADOS, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 26083/124
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-704-711A-18

Query Match 34.4%; Score 45; DB 3; Length 660;
Best Local Similarity 58.8%; Pred. No. 30;
Matches 10; Conservative 0; Mismatches 3; Indels 4; Gaps 1;


```
Oy 1 EFTYMINFGQNYWEH 17
    | | | | | | | |
Db 166 EADIMINFG---WEH 178

RESULT 10
US-08-872-783-3
; Sequence 3, Application US/08872783
; Patent No. 5858717
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: HUMAN FORMIN BINDING PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/872,783
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0317 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 67 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1255023
; US-08-872-783-3.

Query Match 32.8%; Score 43; DB 2; Length 67;
Best Local Similarity 41.2%; Pred. No. 4.9;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Oy 2 FTYMINFGQNYWEHP 18
    | | | | | | | |
Db 51 YTYINYNTGESKWEK 67

RESULT 11
US-08-741-437-3
; Sequence 3, Application US/08741437
; Patent No. 5843665
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: A NOVEL HUMAN PYROPHOSPHATASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/741,437
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0148 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 727225
; US-08-741-437-3

Query Match 32.8%; Score 43; DB 2; Length 114;
Best Local Similarity 38.9%; Pred. No. 8.8;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Oy 4 YMINFGQNYWEHPYQK 21
    | | | | | | | |
Db 8 YIMYGTLPQTWEDPEK 25

RESULT 12
US-09-134-593-3
; Sequence 3, Application US/09134593
; Patent No. 5981232
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: A NOVEL HUMAN PYROPHOSPHATASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,593
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/741,437
; FILING DATE:
```

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0148 US
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 114 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 727225
US-09-134-593-3

Query Match 32.8%; Score 43; DB 2; Length 114;
Best Local Similarity 38.9%; Pred. No. 8.8;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 4 YMINFGGQNYWEHPYQK 21
::: ||| ||| :::
Db 8 YIWNCTLPQWEDPHEK 25

RESULT 13
US-08-773-368-3
Sequence 3, Application US/08773368
Patent No. 5856130
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL HUMAN PATHOGENESIS-RELATED
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/773,368
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy RJ
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0186 US
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 219 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:

LIBRARY: GenBank
CLONE: 847722
US-08-773-368-3

Query Match 32.8%; Score 43; DB 2; Length 219;
Best Local Similarity 38.9%; Pred. No. 18;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 YMINFGGQNYWEHPYQK 21
::: ||| ||| :::
Db 161 FICNYGPGGNYPTWPYKR 178

RESULT 14
US-09-199-887-3
Sequence 3, Application US/09199887
Patent No. 6071874
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL HUMAN PATHOGENESIS-RELATED
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/199,887
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/773,368
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy RJ
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0186 US
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 219 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 847722
US-09-199-887-3

Query Match 32.8%; Score 43; DB 3; Length 219;
Best Local Similarity 38.9%; Pred. No. 18;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 YMINFGGQNYWEHPYQK 21
::: ||| ||| :::
Db 161 FICNYGPGGNYPTWPYKR 178

RESULT 15

US-08-773-368-4
; Sequence 4, Application US/08773368
; Patent No. 5856130
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN PATHOGENESIS-RELATED
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/773,368
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy RJ
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0186 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 266 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1030053
US-08-773-368-4

Query Match 32.8%; Score 43; DB 2; Length 266;
Best Local Similarity 38.9%; Pred. No. 22;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 YMINFGRCQNYWEHPYQK 21
Db 171 FICNYGPGGNYPTWPKR 188

Search completed: January 29, 2002, 10:24:10
Job time: 516 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:26:45 ; Search time 144.96 seconds
(without alignments)
11.561 Million cell updates/sec

Title: US-09-763-397A-19

Perfect score: 131

Sequence: 1 EFTYMINFGQNGYWEHPYQKS 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	131	100.0	622	2 D44986	apical membrane an
2	131	100.0	622	2 A32499	apical membrane an
3	131	100.0	622	2 C44986	apical membrane an
4	126	96.2	622	2 B44986	apical membrane an
5	126	96.2	622	2 A44986	apical membrane an
6	53.5	40.8	631	2 T40189	probable biotin-pr
7	51	38.9	2658	2 A86216	protein T23G18.2 f
8	49	37.4	218	2 T10522	Ribonuclease (EC 3
9	49	37.4	808	2 T39059	probable mannosyl-
10	47	35.9	452	2 C85024	probable polygalac
11	47	35.9	1620	2 E83261	conserved hypothet
12	46	35.1	116	2 T42412	FMRFamide-like pep
13	46	35.1	230	2 A84443	probable ribonucle
14	46	35.1	466	2 E84139	beta-glucosidase B
15	46	35.1	476	1 SGMSV	vitronectin precur
16	46	35.1	502	2 T07141	cytochrome P450 CY
17	45.5	34.7	2100	2 T38128	T7123.15 protein -
18	45	34.4	28	2 PH1911	T-cell receptor al
19	45	34.4	140	2 C49829	T-cell receptor va
20	45	34.4	193	2 T17952	hypothetical prote
21	45	34.4	459	2 JC5139	vitronectin precur
22	45	34.4	465	2 D69785	beta-glucosidase h
23	45	34.4	475	2 A38340	66k glycoprotein p
24	45	34.4	478	1 SGHUIV	vitronectin precur
25	45	34.4	504	2 T70813	hypothetical prote
26	45	34.4	649	2 T27232	hypothetical prote
27	45	34.4	660	1 A28153	gelatinase A (EC 3
28	45	34.4	662	2 A42496	gelatinase A (EC 3
29	45	34.4	662	2 S34780	gelatinase A (EC 3

30 45 34.4 662 2 S70365 gelatinase A (EC 3
31 45 34.4 663 1 S46492 gelatinase A (EC 3
32 44 33.6 104 2 A72570 probable ribosomal
33 44 33.6 196 2 PD0004 self-incompatibili
34 44 33.6 228 2 B86276 hypothetical prote
35 44 33.6 474 2 S77650 probable transposa
36 44 33.6 509 1 S62899 cytochrome P450 (C
37 44 33.6 726 2 S52141 APase - Erwinia a
38 44 33.6 833 2 S62136 CW41 protein - ye
39 44 33.6 1129 2 T25635 hypothetical prote
40 44 33.6 1175 2 T25634 hypothetical prote
41 44 33.6 1216 2 JW0105 synaptojanin 2 alp
42 43.5 33.2 379 2 G70918 hypothetical prote
43 43 32.8 67 2 S64714 formin binding pro
44 43 32.8 80 2 G85867 hypothetical prote
45 43 32.8 219 2 JC4131 glioma pathogenesi

ALIGNMENTS

RESULT 1

D44986

apical membrane antigen 1 - malaria parasite (Plasmodium falciparum) (strain 7c8)
C:Species: Plasmodium falciparum
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000
C:Accession: D44986

R:Thomas, A.W.; Waters, A.P.; Carr, D.

Mol. Biochem. Parasitol. 42, 285-288, 1990

A:Title: Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate a

A:Reference number: A44986; MUID:91101665

A:Accession: D44986

A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: DNA

A:Residues: 1-622 <THO>

A:Cross-references: GB:M34555

C:Keywords: membrane protein; surface antigen

Query Match 100.0%; Score 131; DB 2; Length 622;

Best Local Similarity 100.0%; Pred. No. 6.4e-12;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EFTYMINFGQNGYWEHPYQKS 22

|||||

Db 14 EFTYMINFGQNGYWEHPYQKS 35

RESULT 2

A32499

apical membrane antigen 1 - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 12-Oct-1989 #sequence_revision 12-Oct-1989 #text_change 09-Jun-2000

C:Accession: A32499

R:Peterson, M.G.; Marshall, V.M.; Smythe, J.A.; Crewther, P.E.; Lew, A.; Silva, A.; A

Mol. Cell. Biol. 9, 3151-3154, 1989

A:Title: Integral membrane protein located in the apical complex of Plasmodium falcip

A:Reference number: A32499; MUID:89384584

A:Accession: A32499

A>Status: preliminary

A:Molecule type: DNA; mRNA

A:Residues: 1-622 <PET>

A:Cross-references: GB:M27133; NID:gl60072; PID:gl60073

C:Keywords: membrane protein; surface antigen

Query Match 100.0%; Score 131; DB 2; Length 622;

Best Local Similarity 100.0%; Pred. No. 6.4e-12;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EFTYMINFGQNGYWEHPYQKS 22

|||||

Db 14 EFTYMINFGQNGYWEHPYQKS 35

```
RESULT 3
C44986
apical membrane antigen 1 - malaria parasite (Plasmodium falciparum) (strain FCR 3)
C:Species: Plasmodium falciparum
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000
C:Accession: C44986
R:Thomas, A.W.; Waters, A.P.; Carr, D.
Mol. Biochem. Parasitol. 42, 285-288, 1990
A:Title: Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate anti
A:Accession: A44986
A:Reference number: A44986; MUID:91101665
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: DNA
A:Residues: 1-622 <THO>
A:Cross-references: GB:M34554
C:Keywords: membrane protein; surface antigen

Query Match 96.2%; Score 131; DB 2; Length 622;
Best Local Similarity 100.0%; Pred. No. 6.4e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFTYMINFGRGQNYWEHPYQKS 22
|||||
Db 14 EFTYMINFGRGQNYWEHPYQKS 35
|||||

RESULT 4
B44986
apical membrane antigen 1 - malaria parasite (Plasmodium falciparum) (strain Thai Tn)
C:Species: Plasmodium falciparum
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000
C:Accession: B44986
R:Thomas, A.W.; Waters, A.P.; Carr, D.
Mol. Biochem. Parasitol. 42, 285-288, 1990
A:Title: Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate anti
A:Reference number: A44986; MUID:91101665
A:Accession: B44986
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: DNA
A:Residues: 1-622 <THO>
A:Cross-references: GB:M34553
C:Keywords: membrane protein; surface antigen

Query Match 96.2%; Score 126; DB 2; Length 622;
Best Local Similarity 95.3%; Pred. No. 3.6e-11;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EFTYMINFGRGQNYWEHPYQKS 22
|||||
Db 14 EFTYMINFGRGQNYWEHPYQKS 35
|||||

RESULT 5
A44986
apical membrane antigen 1 - malaria parasite (Plasmodium falciparum) (strain CAMP)
C:Species: Plasmodium falciparum
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000
C:Accession: A44986
R:Thomas, A.W.; Waters, A.P.; Carr, D.
Mol. Biochem. Parasitol. 42, 285-288, 1990
A:Title: Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate anti
A:Reference number: A44986; MUID:91101665
A:Accession: A44986
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: DNA
A:Residues: 1-622 <THO>
A:Cross-references: GB:M34552
C:Keywords: membrane protein; surface antigen
```

```
Query Match 96.2%; Score 126; DB 2; Length 622;
Best Local Similarity 95.5%; Pred. No. 3.6e-11;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EFTYMINFGRGQNYWEHPYQKS 22
|||||
Db 14 EFTYMINFGRGQNYWEHPYQKS 35
|||||

RESULT 6
T40189
probable biotin-protein ligase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T40189
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoeft, A
submitted to the EMBL Data Library, February 1998
A:Reference number: Z21910
A:Accession: T40189
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-631 <WOO>
A:Cross-references: EMBL:Z97992; PIDN:CAB10802.1; GSPDB:GN00067; SPDB:SPBC30D10.07c
A:Experimental source: strain 972h-; cosmid c30D10
C:Genetics:
A:Gene: SPDB:SPBC30D10.07c
A:Map position: 2

Query Match 40.8%; Score 53.5; DB 2; Length 631;
Best Local Similarity 47.8%; Pred. No. 3;
Matches 11; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

QY 2 FTYMINF----GRGQNYWEHPY 19
|| : || : |||
Db 383 FTVLIGNYTAGRGGRQNMWSPY 405
|||||
```

```
RESULT 7
A86216
protein T23G18.2 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: A86216
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Mailli, R.; Marzita
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: A86216
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2658 <STO>
A:Cross-references: GB:AE005172; NID:g6579214; PIDN:AAF18257.1; GSPDB:GN00141
C:Genetics:
A:Gene: T23G18.2
A:Map position: 1
```

```
Query Match 38.9%; Score 51; DB 2; Length 2658;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 MINFGRGQNYWEHP 18
|||
Db 1039 VMNFKRGMNFWEXP 1052
|||||
```

```
RESULT 8
T10522
ribonuclease (EC 3.1.27.-) precursor - Persian tobacco
C:Species: Nicotiana glauca (Persian tobacco)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
C:Accession: T10522
R:Norioka, S.
submitted to the EMBL Data Library, August 1995
A:Description: Molecular cloning and nucleotide sequences of S-RNases and non-S-RNase fr
A:Reference number: 217069
A:Accession: T10522
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-218 <NR>
A:Cross-references: EMBL:D63888
A:Experimental source: style from unopened and mature buds
C:Superfamily: Enterobacter ribonuclease
C:Keywords: hydrolase
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-218/Product: ribonuclease #status predicted <MAT>

Query Match 37.4%; Score 49; DB 2; Length 218;
Best Local Similarity 70.0%; Pred. No. 4.7;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 12 QNYWEHPYOK 21
||| ||| |
Db 104 QNFWEHYNK 113

RESULT 9
T39059
probable mannose-6-phosphate 6-phosphoglucoamido transferase (EC 3.2.1.106) - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 20-Jun-2000
C:Accession: T39059
R:Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1997
A:Reference number: 221824
A:Accession: T39059
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-808 <OLI>
A:Cross-references: EMBL:T98603; PTDN:CAB11295.1; GSPDB:GN000666; SPDB:SPAC6G10.09
A:Experimental source: strain 972h-; cosmid c6G10
C:Genetics:
A:Gene: SPDB:SPAC6G10.09
A:Map position: 1
A:Introns: 80/1
C:Keywords: glycosidase; hydrolase

Query Match 37.4%; Score 49; DB 2; Length 808;
Best Local Similarity 53.8%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 6 INFRGQNYWEHP 18
: || ||| |
Db 711 VYFGTGENYWRGP 723

RESULT 10
C85024
probable polygalacturonidase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C:Accession: C85024
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
```

```
A:Reference number: A85001; MUID:20083488
A:Accession: C85024
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-452 <STO>
A:Cross-references: GB:NC_001268; NID:g7268573; PIDN:CAB80682.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4g01890
A:Map position: 4
C:Superfamily: polygalacturonase

Query Match 35.9%; Score 47; DB 2; Length 452;
Best Local Similarity 70.0%; Pred. No. 20;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 9 GRGQNYWEHP 18
|||||: |
Db 171 GRGQNWDLDP 180

RESULT 11
E83261
conserved hypothetical protein PA3068 [Imported] - Pseudomonas aeruginosa (strain PAO
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: E83261
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: E83261
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1620 <STO>
A:Cross-references: GB:AB004731; GB:AE004091; NID:g9949171; PIDN:AAG06456.1; GSPDB:GN
C:Genetics:
A:Gene: PA3068

Query Match 35.9%; Score 47; DB 2; Length 1620;
Best Local Similarity 36.8%; Pred. No. 79;
Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 1 EFTYMINFRGQNYWEHPY 19
|| | : || | :
Db 597 EFPYRLRHQNGREYWIHDF 615

RESULT 12
T42412
FMRFamide-like peptide 5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T42412
R:Li, C.; Nelson, L.S.; Memmott, J.M.
submitted to the EMBL Data Library, January 1998
A:Reference number: 222162
A:Accession: T42412
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-116 <LIC>
A:Cross-references: EMBL:AF042391; PIDN:AAC08942.1
C:Genetics:
A:Gene: flp-5

Query Match 35.1%; Score 46; DB 2; Length 116;
Best Local Similarity 57.1%; Pred. No. 6.8;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

QY 6 INFGRGQNYWEHPY 19

I : I I I I

Db 83 IFRGSRNTWEDGY 96

RESULT 13

A84443

probable ribonuclease, RNS1 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001

C:Accession: A84443

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487

A:Accession: A84443

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-230 <STO>

A:Cross-references: GB:AE002093; NID:g3461823; PIDN:AAC32917.1; GSPDB:GN00139

C:Genetics:

A:Gene: AC2g02990

A:Map position: 2

C:Superfamily: Enterobacter ribonuclease

Query Match

Best Local Similarity

Matches

Score 35.1%; DB 2; Length 230;

Indels 3; Mismatches 4; Gaps 0;

Conservative 6;

QY 9 GRGQNYWEHPYQK 21

I : I I I I

Db 110 GSGEAFWEHEWK 122

RESULT 14

F84139

beta-glucosidase BH3918 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000

C:Accession: F84139

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20263314

A:Accession: F84139

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-466 <STO>

A:Cross-references: GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BAB07637.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH3918

C:Superfamily: Agrobacterium beta-glucosidase

Query Match

Best Local Similarity

Matches

Score 35.1%; DB 2; Length 466;

Indels 2; Mismatches 4; Gaps 0;

Conservative 6;

QY 9 GRGQNYWEHPYQ 20

I : I I I I

Db 33 GKGNINWDHWE 44

RESULT 15

SGMSV

vitronectin precursor - mouse

N:Alternate names: complement protein S

N:Centains: somatomedin B

C:Species: Mus musculus (house mouse)

C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 22-Jun-1999

C:Accession: S19894; JT0662; S22224; A56398

R:Ehrlich, H.J.; Richter, B.; von der Ahe, D.; Preissner, K.T.

submitted to the EMBL Data Library, November 1991

A:Description: Molecular cloning and expression in E. coli of mouse vitronectin.

A:Reference number: S19894

A:Accession: S19894

A:Molecule type: mRNA

A:Residues: 1-476 <EHR>

A:Cross-references: EMBL:X63003

R:Seiffert, D.; Poenninger, J.; Binder, B.R.

Gene 134, 303-304, 1993

A:Title: Organization of the gene encoding mouse vitronectin.

A:Reference number: JT0662; MUID:94085797

A:Accession: JT0662

A:Molecule type: DNA

A:Residues: 1-255, 'A', 256-434, 'S', 435-476 <SEI>

A:Cross-references: GB:X72091; NID:9441465; PIDN:CAAS0981.1; PID:g441466

R:Nakashima, N.; Miyazaki, K.; Ishikawa, M.; Yatohgo, T.; Ogawa, H.; Uchibori, H.; Ma

Biochim. Biophys. Acta 1120, 1-10, 1992

A:Title: Vitronectin diversity in evolution but uniformity in ligand binding and size

A:Reference number: S21768; MUID:92207982

A:Accession: S22224

A:Status: preliminary

A:Molecule type: protein

A:Residues: 20, 'XXXF', 25-27, 'X', 29-34, 'X', 36-37, 'X', 39, 'X', 41, 'X', 43, 'X' <NAK>

R:Seiffert, D.; Keeton, M.; Eguchi, Y.; Sawdey, M.; Loskutoff, D.J.

Proc. Natl. Acad. Sci. U.S.A. 88, 9402-9406, 1991

A:Title: Detection of vitronectin mRNA in tissues and cells of the mouse.

A:Reference number: A56398; MUID:92052101

A:Accession: A56398

A:Molecule type: mRNA

A:Residues: 1-255, 'A', 256-382, 385-414, 'K', 416-434, 'S', 435-476 <SE2>

A:Cross-references: GB:M77123; NID:g202371; PIDN:AAA40558.1; PID:g202372

C:Genetics:

A:Introns: 22/1; 62/1; 176/1; 222/3; 273/1; 324/1; 439/1

C:Superfamily: vitronectin; hemopexin repeat homology; somatomedin B homology

C:Keywords: cell adhesion; glycoprotein; heparin binding; phosphoprotein; sulfoprotei

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-476/Product: vitronectin #status predicted <MAT>

F:20-63/Domain: somatomedin B #status predicted <SOM>

F:20-62/Region: cell attachment (R-G-D) motif

F:84-66/Region: hemopexin repeat homology #status atypical <PX1>

F:150-285/Domain: hemopexin repeat homology #status atypical <PX2>

F:286-471/Domain: hemopexin repeat homology #status atypical <PX2>

F:365-398/Region: heparin binding #status predicted

F:75-78/Binding site: sulfate (Tyr) (covalent) #status predicted

F:86-168,241/Binding site: carboxylate (Asn) (covalent) #status predicted

F:291-430/Disulfide bonds: #status predicted

F:397/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match

Best Local Similarity

Matches

Score 35.1%; DB 1; Length 476;

Indels 4; Mismatches 6; Gaps 0;

Conservative 6;

QY 6 INFGRGQNYWEHPYQK 21

I : I I I I

Db 269 VYFFKQKQWYEFQ 284

Search completed: January 29, 2002, 10:26:46

Job time: 657 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:13:47 ; Search time 80.65 seconds
(without alignments)
10.002 Million cell updates/sec

Title: US-09-763-397a-19

Perfect score: 131
Sequence: 1 EFTYMINFGRGQNYWEHPYQKS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	131	100.0	622	1 AMAL_PLAF8	P50492 plasmodium
2	131	100.0	622	1 AMAL_PLAFF	P22621 plasmodium
3	131	100.0	622	1 AMAL_PLAFG	P50490 plasmodium
4	126	96.2	622	1 AMAL_PLAFH	P50489 plasmodium
5	126	96.2	622	1 AMAL_PLAFH	P50491 plasmodium
6	49	37.4	808	1 GCSL_SCHPO	O14255 schizosacch
7	46.5	35.5	321	1 ABNA_ASPNG	P42256 aspergillus
8	46	35.1	230	1 RNS1_ARATH	P42813 arabidopsis
9	46	35.1	478	1 VTNC_MOUSE	P29788 mus musculus
10	46	35.1	502	1 C932_SOYBN	P42799 glycine max
11	45	34.4	459	1 VTNC_PIG	P48819 sus scrofa
12	45	34.4	475	1 VTNC_RABIT	P22458 oryctolagus
13	45	34.4	478	1 VTNC_HUMAN	P04004 homo sapien
14	45	34.4	660	1 MM02_HUMAN	P08253 homo sapien
15	45	34.4	662	1 MM02_MOUSE	P33434 mus musculus
16	45	34.4	662	1 MM02_RABIT	P50757 oryctolagus
17	45	34.4	662	1 MM02_RAT	P33436 rattus norv
18	45	34.4	663	1 MM02_CHICK	Q90611 gallus gall
19	44	33.6	104	1 RS10_AERPE	Q9yav2 aeropyrum p
20	44	33.6	186	1 ARL6_HUMAN	Q9h0f7 homo sapien
21	44	33.6	218	1 RNS7_NICAL	Q40381 nicotiana a
22	44	33.6	509	1 C931_SOYBN	O42798 glycine max
23	44	33.6	726	1 AMSA_ERWAM	O46631 erwinia amy
24	44	33.6	833	1 CW41_YEAST	P33008 saccharomyc
25	44	33.6	1248	1 SYJ2_RAT	Q55207 rattus norv
26	44	33.6	2168	1 POLG_PEV9U	O41174 p genome po
27	43	32.8	186	1 ARL6_MOUSE	O88848 mus musculus
28	43	32.8	266	1 GLIP_HUMAN	P48060 homo sapien
29	43	32.8	297	1 YX01_CAEEL	Q11108 caenorhabdi
30	43	32.8	420	1 YN57_YEAST	Q33694 saccharomyc
31	43	32.8	1699	1 POLN_LORDV	P54634 lordsdale v
32	43	32.8	1782	1 VIT_BOMMO	Q27309 bombyx mori
33	43	32.8	2298	1 C215_HUMAN	Q9y3r5 homo sapien

ALIGNMENTS

RESULT 1	AMAL_PLAF8	STANDARD:	PRT:	622 AA.
ID	AMAL_PLAF8			
AC	P50492:			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).			
GN	AMA-1 OR PF83.			
OS	Plasmodium falciparum (isolate 7G8).			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
OX	NCBI_TaxID=57266;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91101665; PubMed=2270110;			
RA	Thomas A.W., Waters A.P., Carr D.;			
RT	"Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate antigen of Plasmodium falciparum.";			
RL	Mol. Biochem. Parasitol. 42:285-287(1990).			
CC	-1- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.			
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	-1- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO PK66 FROM P.KNOWLES.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M58548; AAA29721.1; ..			
DR	InterPro: IPR003298; Apmem_Ag1.			
DR	Pfam; PF02430; AMA-1; 1.			
KW	Malaria; Signal; Transmembrane; Antigen; Glycoprotein.			
FT	SIGNAL 1 24 POTENTIAL.			
FT	CHAIN 25 622 APICAL MEMBRANE ANTIGEN 1.			
FT	DOMAIN 25 546 EXTRACELLULAR (POTENTIAL).			
FT	TRANSMEM 547 567 POTENTIAL.			
FT	DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).			
FT	CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	CARBOHYD 421 421 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	CARBOHYD 422 422 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	CARBOHYD 499 499 N-LINKED (GLCNAC. . .) (POTENTIAL).			
SQ	SEQUENCE 622 AA; 71990 MW; 1C9C8715D8E2915F CRC64;			

Query Match 100.0% Score 131; DB 1; Length 622;
Best Local Similarity 100.0%; Pred. No. 5.9e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 EFTYMINFGRGQNYWEHPYQKS 22
|||||

```
Db 14 EFTYMINFGRGQNYWHPYOKS 35

RESULT 2
AMAL_PLAFC
ID AMAL_PLAFC STANDARD; PRT; 622 AA.
AC P50490;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
GN AMA-1 OR PF83.
OS Plasmodium falciparum (isolate FC27 / Papua New Guinea).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5837;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91101665; PubMed=2270110;
RA Thomas A.W., Waters A.P., Carr D.;
RT "Analysis of variation in PF83, an erythrocytic merozoite vaccine
candidate antigen of Plasmodium falciparum."
RL Mol. Biochem. Parasitol. 42:285-287(1990).
CC -!- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO
PK66 FROM P.KNOWLES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M58546; AAA29475.1; -
DR InterPro: IPR003298; Apmem_Agl.
DR Pfam; PF02430; AMA-1; 1.
DR Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
KW SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 622 APICAL MEMBRANE ANTIGEN 1.
FT DOMAIN 25 546 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 547 567 POTENTIAL.
FT DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 162 162 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 422 422 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 503 503 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 622 AA; 72009 MW; 7D41335E249FA18F CRC64;

Query Match 100.0%; Score 131; DB 1; Length 622;
Best Local Similarity 100.0%; Pred. No. 5.9e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFTYMINFGRGQNYWHPYOKS 22
Db 14 EFTYMINFGRGQNYWHPYOKS 35

RESULT 3
AMAL_PLAFC
ID AMAL_PLAFC STANDARD; PRT; 622 AA.
AC P50489;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
GN AMA-1 OR PF83.
OS Plasmodium falciparum (isolate Camp / Malaysia).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5835;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91101665; PubMed=2270110;

Query Match 100.0%; Score 131; DB 1; Length 622;
Best Local Similarity 100.0%; Pred. No. 5.9e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFTYMINFGRGQNYWHPYOKS 22
Db 14 EFTYMINFGRGQNYWHPYOKS 35
```

RA Thomas A.W., Waters A.P., Carr D.;
RT "Analysis of variation in PF83, an erythrocytic merozoite vaccine
RL candidate antigen of Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 42:285-287(1990).
CC -!- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO
CC PK66 FROM P.KNOWLES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M58545; AAA29718.1; -;
DR InterPro; IPR003298; Apmem_Ag1.
DR Pfam; PF02430; AMA-1; 1.
KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 622 APICAL MEMBRANE ANTIGEN 1.
FT DOMAIN 25 546 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 547 567 POTENTIAL.
FT DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 162 162 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 422 422 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 622 AA; 71943 MW; 26CE8CF76D07C637 CRC64;

Query Match 96.2%; Score 126; DB 1; Length 622;
Best Local Similarity 95.5%; Pred. No. 3.2e-11;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EFTYMINFGRGQNYWEHPYQKS 22
Db 14 EFTYMINFGRGQNYWEHPYQNS 35
|||||
RESULT 5
ANAL_PLAFH STANDARD; PRT; 622 AA.
ID AMAL_PLAFH
AC P50491;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
GN AMA-1 OR PF83.
OS Plasmodium falciparum (isolate thn / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=70151;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91101665; PubMed=2270110;
RA Thomas A.W., Waters A.P., Carr D.;
RT "Analysis of variation in PF83, an erythrocytic merozoite vaccine
RL candidate antigen of Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 42:285-287(1990).
CC -!- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO
CC PK66 FROM P.KNOWLES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M58547; AAA29720.1; -;
DR InterPro; IPR003298; Apmem_Ag1.
DR Pfam; PF02430; AMA-1; 1.
KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 622 APICAL MEMBRANE ANTIGEN 1.
FT DOMAIN 25 546 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 547 567 POTENTIAL.
FT DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 422 422 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 622 AA; 71989 MW; 1FDFA53593C94CC5 CRC64;

Query Match 96.2%; Score 126; DB 1; Length 622;
Best Local Similarity 95.5%; Pred. No. 3.2e-11;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EFTYMINFGRGQNYWEHPYQKS 22
Db 14 EFTYMINFGRGQNYWEHPYQNS 35
|||||
RESULT 6
GCSI_SCHPO STANDARD; PRT; 808 AA.
ID GCSI_SCHPO
AC O14255;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PROBABLE MANNOSYL-OLIGOSACCHARIDE GLUCOSIDASE (EC 3.2.1.106)
DE (PROCESSING A-GLUCOSIDASE I).
GN SPAC6G10.09.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Oliver K., Harris D., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CLEAVES THE DISTAL ALPHA 1,2-LINKED GLUCOSE RESIDUE FROM
CC THE GLC(3)MAN(9)GLCNAC(2) OLIGOSACCHARIDE PRECURSOR HIGHLY
CC SPECIFICALLY (BY SIMILARITY)
CC -!- CATALYTIC ACTIVITY: EXOHYDROLYSIS OF THE NON-REDUCING TERMINAL
CC GLUCOSE RESIDUE IN THE MANNOSYL-OLIGOSACCHARIDE
CC GLC(3)MAN(9)GLCNAC(2).
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ENDOPLASMIC
CC RETICULUM (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO FAMILY 63 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z98603; CAB11295.1; -;
KW Hypothetical protein; Hydrolase; Glycosidase; Glycoprotein;
KW Transmembrane; Signal-anchor; Endoplasmic reticulum.
FT DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 12 31 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT

FT DOMAIN 32 808 LUMINAL (POTENTIAL).
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 808 AA; 92889 MW; 31871729E6F95254 CRC64;

Query Match 37.4%; Score 49; DB 1; Length 808;
 Best Local Similarity 53.8%; Pred. No. 9.7;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 INFGRCQNYWEHP 18

Db 711 VYFGTGENYWRGP 723

RESULT 7
 ABNA_ASPNG STANDARD; PRT; 321 AA.
 AC P42256;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ARABINAN ENDO-1,5-ALPHA-L-ARABINOSIDASE A PRECURSOR (EC 3.2.1.99)
 DE (ENDO-1,5-ALPHA-L-ARABINANASE A) (ABN A).
 GN ABNA.
 OS Aspergillus niger.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5061;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-32 AND 125-139.
 RC STRAIN=CBS 120.49 / N400;
 RX MEDLINE=94128348; PubMed=7764386;
 RA Flippin M.J.A., Panneman H., van der Veen P., Visser J.,
 de Graaf L.H.;
 RT "Molecular cloning, expression and structure of the endo-1,5-alpha-L-
 arabinase gene of Aspergillus niger."
 RL Appl. Microbiol. Biotechnol. 40:318-326(1993).
 CC -!- FUNCTION: ITS PREFERRED SUBSTRATE IS LINEAR 1,5-ALPHA-L-ARABINAN.
 CC THE ENZYME ACTIVITY IS PROGRESSIVELY REDUCED AS 1,5-ALPHA-CHAINS
 CC BECOME SHORTER OR MORE HIGHLY SUBSTITUTED.
 CC -!- CATALYTIC ACTIVITY: ENDOLYOLYSIS OF 1,5-ALPHA-L-
 ARABINOFURANOSIDIC LINKAGES IN 1,5-ARABINANS.
 CC -!- PATHWAY: INVOLVED IN DEGRADATION OF THE PLANT CELL WALL
 CC POLYSACCHARIDE L-ARABINAN.
 CC -!- INDUCTION: BY L-ARABAN, ARABINOGLACTAN AND L-ARABITOL.
 CC -!- SIMILARITY: BELONGS TO FAMILY 43 OF GLYCOSYL HYDROLASES.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

 EMBL: L23430; AAA32682.1;
 KW Hydrolase; Glycosidase; Glycoprotein; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 321
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SEQUENCE 321 AA; 34454 MW; 3087DD7D0592A0B20 CRC64;
 FT A.
 FT ARABINAN ENDO-1,5-ALPHA-L-ARABINOSIDASE

Query Match 35.5%; Score 46.5; DB 1; Length 321;
 Best Local Similarity 50.0%; Pred. No. 9.2;
 Matches 9; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

QY 3 TYMIFGRQNYWEHPYQ 20

Db 158 TYVINFG---SEWDDIYQ 172

• . . .

RESULT 8

RNSI_ARATH

ID RNSI_ARATH STANDARD; PRT; 230 AA.

AC P42813;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE RIBONUCLEASE 1 PRECURSOR (EC 3.1.27.1).

GN RNSI OR AT2G02990 OR T17M13.16.

OS Arabidopsis thaliana (Mouse-ear cross).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RX MEDLINE=95093473; PubMed=8000425;

RA Bariola P.A., Howard C.J., Taylor C.B., Verburg M.T., Jaglan V.D.,

Green P.J.;

RT "The Arabidopsis ribonuclease gene RNS1 is tightly controlled in

response to phosphate limitation."

RL Plant J. 6:673-685(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RX MEDLINE=20083487; PubMed=10617197;

RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,

Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,

Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,

Moffat K.S., Cronin L.A., Shen M., Vanaken S.E., Umayam L.,

Taiton L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,

Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,

RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,

Venter J.C.;

RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis

thaliana."

RL Nature 402:761-768(1999).

RN [3]

RP SEQUENCE OF 23-41.

RA Robertson D., Mitchell G.P., Gilroy J.S., Gerrish C., Bolwell G.P.,

Slabas A.R.;

RL Submitted (JAN-1997) to the SWISS-PROT data bank.

CC -!- FUNCTION: MAY REMOILIZE PHOSPHATE, PARTICULARLY WHEN CELLS

SENESCE OR WHEN PHOSPHATE BECOMES LIMITING.

CC -!- CATALYTIC ACTIVITY: TWO-STAGE ENDONUCLEOLYTIC CLEAVAGE TO 3'-

PHOSPHOMONONUCLEOTIDES AND 3'-PHOSPHOLIGONUCLEOTIDES WITH

2',3'-CYCLIC PHOSPHATE INTERMEDIATES.

CC -!- INDUCTION: BY PHOSPHATE-STARVATION.

CC -!- SIMILARITY: BELONGS TO THE RNASE T2 FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

 EMBL: U05206; AAC48925.1;
 DR EMBL; AC004138; AAC32917.1; -.
 DR HSSP; P08056; IBOL.
 DR InterPro; IPR001568; RNase_T2.
 DR Pfam; PF00445; ribonuclease_T2; 1.
 DR PROSITE; PS00530; RNASE_T2.1; 1.
 DR PROSITE; PS00531; RNASE_T2.2; 1.
 KW Hydrolase; Nuclease; Endonuclease; Signal; Multigene family.

FT SIGNAL 1 22

FT CHAIN 23 230 RIBONUCLEASE 1.

FT DISULFID 80 126 BY SIMILARITY.

FT DISULFID 186 221 BY SIMILARITY.

FT ACT_SITE 65 65 BY SIMILARITY.

FT ACT_SITE 119 119 BY SIMILARITY.

FT ACT_SITE 123 123 BY SIMILARITY.
FT CONFLICT 39 41 WPG -> GXP (IN REF. 3).
SQ SEQUENCE 230 AA; 25396 MW; DFD132D39F02505A CRC64;

Query Match 35.1%; Score 46; DB 1; Length 230;
Best Local Similarity 46.2%; Pred. No. 7.8;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 GRGONYWEHPYQK 21
I I : : I I I : : I
Db 110 GSGEAFWEHEWK 122

RESULT 9
VTNC_MOUSE STANDARD; PRT; 478 AA.
AC P29788;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE VITRONECTIN PRECURSOR (SERUM SPREADING FACTOR) (S-PROTEIN).
GN VTN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE=Liver;
RX MEDLINE=94085797; PubMed=7505250;
RA Seifert D., Keeton M., Eguchi Y., Sawdey M., Loskutoff D.J.;
RT "Detection of vitronectin mRNA in tissues and cells of the mouse.";
RL Gene 134:303-304(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92052101; PubMed=1719529;
RA Seifert D., Keeton M., Eguchi Y., Sawdey M., Loskutoff D.J.;
RT "Detection of vitronectin mRNA in tissues and cells of the mouse.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9402-9406(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE=Liver;
RA Ehrlich H.J., Richter B., von der Ahe D., Preissner K.T.;
RT "Primary structure of vitronectins and homology with other proteins.";
RL (In) Preissner K.T., Rosenblatt S., Kost C., Wegerhoff J.,
RL Mosher D.F. (eds.);
RL Biology of vitronectins, pp.1-1, Elsevier, Amsterdam (1996).
CC -!- FUNCTION: VITRONECTIN IS A CELL ADHESION AND SPREADING FACTOR
CC FOUND IN SERUM AND TISSUES. VITRONECTIN INTERACT WITH
CC GLYCOSAMINOGLYCANS AND PROTEOGLYCANS. IS RECOGNIZED BY CERTAIN
CC MEMBERS OF THE INTEGRIN FAMILY AND SERVES AS A CELL-TO-SUBSTRATE
CC ADHESION MOLECULE. INHIBITOR OF THE MEMBRANE-DAMAGING EFFECT OF
CC THE TERMINAL CYTOLYTIC COMPLEMENT PATHWAY.
CC -!- FUNCTION: SOMATOMEDIN B IS A GROWTH HORMONE-DEPENDENT SERUM FACTOR
CC WITH PROTEASE-INHIBITING ACTIVITY.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- TISSUE SPECIFICITY: PLASMA.
CC -!- PTM: SULFATED ON 2 TYROSINE RESIDUES (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 2 HEMOPEXIN-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 SOMATOMEDIN-B TYPE DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
or send an email to license@isb-sib.ch).

EMBL; X72091; CAA50981.1; -;
EMBL; M71123; AAA40558.1; -;

DR EMBL; X63003; CAA44732.1; -;
DR PIR; S19894; SGMV.
DR HSP; P45452; LPEX.
DR MGD; MGI:98940; Vtn.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR001212; Somatomedin_B.
DR Pfam; PF00045; hemopexin; 4.
DR Pfam; PF01033; Somatomedin_B; 1.
DR PRINTS; PR00023; SOMATOMEDINB..
DR SMART; SM00120; HX; 4.
DR SMART; SM00201; SO; 1.
DR PROSITE; PS00024; HEMOPEXIN; 2.
DR PROSITE; PS00524; SOMATOMEDIN_B; 1.
KW Heparin-binding; Cell adhesion; Glycoprotein; Sulfation; Signal;
KW Phosphorylation.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 478 VITRONECTIN.
FT DOMAIN 20 63 SOMATOMEDIN-B LIKE.
FT DOMAIN 150 286 HEMOPEXIN-LIKE 1.
FT DOMAIN 287 478 HEMOPEXIN-LIKE 2.
FT SITE 64 66 CELL ATTACHMENT SITE.
FT MOD_RES 75 75 SULFATION (POTENTIAL).
FT MOD_RES 78 78 SULFATION (POTENTIAL).
FT MOD_RES 80 80 SULFATION (POTENTIAL).
FT MOD_RES 278 278 SULFATION (POTENTIAL).
FT MOD_RES 281 281 SULFATION (POTENTIAL).
FT MOD_RES 416 416 SULFATION (POTENTIAL).
FT MOD_RES 419 419 SULFATION (POTENTIAL).
FT MOD_RES 421 421 SULFATION (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 292 431 BY SIMILARITY.
FT BINDING 366 399 HEPARIN (BY SIMILARITY).
FT MOD_RES 398 398 PHOSPHORYLATION (BY CAPK)
FT CONFLICT 383 384 (BY SIMILARITY).
FT CONFLICT 416 416 MISSING (IN REF. 2).
FT CONFLICT 478 AA; 54849 MW; EB0C772F8BD6A166 CRC64;
SQ SEQUENCE 478 AA; 54849 MW; EB0C772F8BD6A166 CRC64;

Query Match 35.1%; Score 46; DB 1; Length 478;
Best Local Similarity 37.5%; Pred. No. 16;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 6 INFGRGQNYWEHPYQK 21
I I : : I I I : : I
Db 270 VYFFKQKQYWEYEQ 285

RESULT 10
C932_SOYBN STANDARD; PRT; 502 AA.
AC C932_SOYBN
AC Q42799;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CYTOCHROME P450 93A2 (EC 1.14.-.-).
GN CYP93A2.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RA Suzuki G., Ohta H., Kato T., Shibata D., Masuda T., Takamiya K.-I.;
RT "Molecular cloning of a cDNA encoding cytochrome P450 CYP93A2 from
RT soybean suspension-cultured cells."
RL (In) Plant Gene Register PGR97-087.
CC -!- INDUCTION: BY METHYL JASMONATE (MEJA).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: D86351; BAA13076.1; -
 DR InterPro: IPR001128; Cyt_P450.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Membrane; Heme.
 FT BINDING 440 440 HEME (BY SIMILARITY).
 SQ SEQUENCE 502 AA; 57176 MW; D85C419D0698C77C CRC64;

Query Match 35.1%; Score 46; DB 1; Length 502;
 Best Local Similarity 50.0%; Pred. No. 17;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 9 GRGQNYWEHPYQ 20

DB 396 GRDPNHNWPF 407

RESULT 11

VTNC_PIG STANDARD; PRT; 459 AA.

AC P48819;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE VITRONECTIN PRECURSOR (SERUM SPREADING FACTOR) (S-PROTEIN).
 GN VTN.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 ON NCBI_TaxID=9823;
 RX SEQUENCE FROM N.A.
 RP MEDLINE=97137525; PubMed=8982862;
 RA Yoneda A., Kojima K., Matsumoto I., Yamamoto K., Ogawa H.;
 RT "Porcine vitronectin, the most compact form of single-chain
 RT vitronectin: the smallest molecular mass among vitronectins was
 RT ascribed to deletion and substitution of base pairs, and proteolytic
 RT trimming of the peptide".
 RL J. Biochem. 120:954-960(1996).

CC -1- FUNCTION: VITRONECTIN IS A CELL ADHESION AND SPREADING FACTOR
 CC FOUND IN SERUM AND TISSUES. VITRONECTIN INTERACT WITH
 CC GLYCOSAMINOGLYCANS AND PROTEOGLYCANS. IS RECOGNIZED BY CERTAIN
 CC MEMBERS OF THE INTEGRIN FAMILY AND SERVES AS A CELL-TO-SUBSTRATE
 CC ADHESION MOLECULE. INHIBITOR OF THE MEMBRANE-DAMAGING EFFECT OF
 CC THE TERMINAL CYTOLYTIC COMPLEMENT PATHWAY (BY SIMILARITY).
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: SULFATED ON 2 TYROSINE RESIDUES (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 2 HEMOPEXIN-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SOMATOMEDIN-B TYPE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: D63145; BAA09630.1; -

DR HSSP: P45452; IPEX.

DR GlycoSuiteDB; P48819; -

DR InterPro: IPR000585; Hemopexin.
 DR InterPro: IPR001212; Somatomedin_B.
 DR Pfam: PF00045; hemopexin; 4.
 DR Pfam: PF01033; Somatomedin_B; 1.
 DR PRINTS: PR00022; SOMATOMEDINB.
 DR SMART: SM00120; HX; 4.
 DR SMART: SM00201; SO; 1.
 DR PROSITE: PS00024; HEMOPEXIN; 1.
 DR PROSITE: PS00524; SOMATOMEDIN_B; 1.
 KW Heparin-binding; Cell adhesion; Glycoprotein; Sulfation; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 459 VITRONECTIN.
 FT DOMAIN 20 63 SOMATOMEDIN-B LIKE.
 FT DOMAIN 128 264 HEMOPEXIN-LIKE 1.
 FT DOMAIN 265 459 HEMOPEXIN-LIKE 2.
 FT SITE 64 66 CELL ATTACHMENT SITE (POTENTIAL).
 FT MOD_RES 75 75 SULFATION (POTENTIAL).
 FT MOD_RES 78 78 SULFATION (POTENTIAL).
 FT MOD_RES 80 80 SULFATION (POTENTIAL).
 FT MOD_RES 398 398 SULFATION (POTENTIAL).
 FT MOD_RES 401 401 SULFATION (POTENTIAL).
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 459 AA; 52572 MW; 630E134B0DC706BE CRC64;

Query Match 34.4%; Score 45; DB 1; Length 459;

Best Local Similarity 37.5%; Pred. No. 22;

Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 6 INFGRGQNYWEHPYQ 21

DB 248 VYFFKGQYWEYVFOQ 263

RESULT 12

VTNC_RABIT

ID VTNC_RABIT STANDARD; PRT; 475 AA.

AC P22458;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE VITRONECTIN PRECURSOR (SERUM SPREADING FACTOR) (S-PROTEIN)
 GN VTN.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 ON NCBI_TaxID=9986;
 RX SEQUENCE FROM N.A.; AND SEQUENCE OF 1-24.
 RP MEDLINE=91065939; PubMed=1701177;
 RA Sato R., Komine Y., Imanaka T., Takano T.;
 RT "Monoclonal antibody EMRIA/212D recognizing site of deposition of
 RT extracellular lipid in atherosclerosis. Isolation and
 RT characterization of a cDNA clone for the antigen".
 RL J. Biol. Chem. 265:21232-21236(1990).

CC -1- FUNCTION: VITRONECTIN IS A CELL ADHESION AND SPREADING FACTOR
 CC FOUND IN SERUM AND TISSUES. VITRONECTIN INTERACT WITH
 CC GLYCOSAMINOGLYCANS AND PROTEOGLYCANS. IS RECOGNIZED BY CERTAIN
 CC MEMBERS OF THE INTEGRIN FAMILY AND SERVES AS A CELL-TO-SUBSTRATE
 CC ADHESION MOLECULE. INHIBITOR OF THE MEMBRANE-DAMAGING EFFECT OF
 CC THE TERMINAL CYTOLYTIC COMPLEMENT PATHWAY.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: SULFATED ON 2 TYROSINE RESIDUES (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 2 HEMOPEXIN-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SOMATOMEDIN-B TYPE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M55442; AAA31258.1; -
PIR; A38340; A38340.
HSSP; P45452; IPEX.
InterPro: IPR000585; Hemopexin.
InterPro: IPR001212; Somatomedin_B.
Pfam; PF00045; hemopexin; 4.
Pfam; PF01033; Somatomedin_B; 1.
PRINTS; PR00022; SOMATOMEDINB.
SMART; SM00120; HX; 4.
SMART; SM00201; SO; 1.
PROSITE; PS00024; HEMOPEXIN; 2.
PROSITE; PS00524; SOMATOMEDIN_B; 1.
Heparin-binding; Cell adhesion; Glycoprotein; Sulfation; Signal.
FT SIGNAL 1 19
FT CHAIN 20 475
FT DOMAIN 20 63
FT DOMAIN 150 287
FT DOMAIN 288 475
FT DOMAIN 366 392
FT SITE 64 66
FT MOD_RES 75 75
FT MOD_RES 78 78
FT MOD_RES 80 80
FT MOD_RES 279 279
FT MOD_RES 282 282
FT CARBOHYD 87 87
FT CARBOHYD 169 169
FT CARBOHYD 242 242
SQ SEQUENCE 475 AA; 53943 MW; DSD1F31B8C2FA12D CRC64;
Query Match 34.4%; Score 45; DB 1; Length 475;
Best Local Similarity 37.5%; Pred. No. 22;
Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
Qy 6 INFGKQNYWEHPYQK 21
Db 271 VYFFKGDYWEYFQQ 286
RESULT 13
VTNC_HUMAN
ID VTNC_HUMAN STANDARD; PRT; 478 AA.
AC P04004; P01141;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE VITRONECTIN PRECURSOR (SERUM SPREADING FACTOR) (S-PROTEIN) [CONTAINS:
DE SOMATOMEDIN B].
GN VTN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86030229; PubMed=2414098;
RA Suzuki S., Oldberg A., Hayman E.G., Pierschbacher M.D., Ruoslahti E.;
RT "Complete amino acid sequence of human vitronectin deduced from cDNA.
RT Similarity of cell attachment sites in vitronectin and fibronectin.";
RL EMBO J. 4:2519-2524(1985).
RN [2]
RP REVISIONS.
RA Suzuki S., Oldberg A., Hayman E.G., Pierschbacher M.D., Ruoslahti E.;
RL Submitted (JUN-1986) to the PIR data bank.
RP [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86135941; PubMed=3004934;

RA Jenne D., Stanley K.K.;
RT "Molecular cloning of S-protein, a link between complement,
RT coagulation and cell-substrate adhesion.";
RL EMBO J. 4:3153-3157(1985).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=88107592; PubMed=2447940;
RA Jenne D., Stanley K.K.;
RT "Nucleotide sequence and organization of the human S-protein gene:
RT repeating peptide motifs in the 'pexin' family and a model for their
RT evolution.";
RL Biochemistry 26:6735-6742(1987).
RN [5]
RP SEQUENCE OF 20-63.
RX MEDLINE=78127267; PubMed=631332;
RA Fryklund L., Sievertsson H.;
RT "Primary structure of somatomedin B: a growth hormone-dependent serum
RT factor with protease inhibiting activity.";
RL FEBS Lett. 87:55-60(1978).
RN [6]
RP SEQUENCE OF 399-413.
RC TISSUE=Plasma;
RA Hochstrasser D.F., Frutiger S., Paquet N., Bairoch A., Ravier F.,
RA Pasquali C., Sanchez J.-C., Tissot J.-D., Bjellqvist B., Vargas R.,
RA Appel R.D., Hughes G.J.;
RL Submitted (JUN-1992) to the SWISS-PROT data bank.
RN [7]
RP SULFATION.
RX MEDLINE=90060125; PubMed=2479556;
RA Jenne D., Hille A., Stanley K.K., Huttner W.B.;
RT "Sulfation of two tyrosine-residues in human complement S-protein
RT (vitronectin).";
RL Eur. J. Biochem. 185:391-395(1989).
RN [8]
RP PHOSPHORYLATION.
RX MEDLINE=90353578; PubMed=1696913;
RA Chain D., Korc-Grodzicki B., Kreizman T., Shaltiel S.;
RT "The phosphorylation of the two-chain form of vitronectin by protein
RT kinase A is heparin dependent.";
RL FEBS Lett. 269:221-225(1990).
CC -!- FUNCTION: VITRONECTIN IS A CELL ADHESION AND SPREADING FACTOR
CC FOUND IN SERUM AND TISSUES. VITRONECTIN INTERACT WITH
CC GLYCOSAMINOGLYCANS AND PROTEOGLYCANS. IS RECOGNIZED BY CERTAIN
CC MEMBERS OF THE INTEGRIN FAMILY AND SERVES AS A CELL-TO-SUBSTRATE
CC ADHESION MOLECULE. INHIBITOR OF THE MEMBRANE-DAMAGING EFFECT OF
CC THE TERMINAL CYTOLYTIC COMPLEMENT PATHWAY.
CC -!- FUNCTION: SOMATOMEDIN B IS A GROWTH HORMONE-DEPENDENT SERUM FACTOR
CC WITH PROTEASE-INHIBITING ACTIVITY.
CC -!- SUBUNIT: EXISTS IN TWO FORMS: A SINGLE CHAIN 75 KDA FORM (V75) AND
CC WHICH ARE HELD TOGETHER BY A DISULFIDE BOND.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- TISSUE SPECIFICITY: PLASMA.
CC -!- PTM: SULFATED ON 2 TYROSINE RESIDUES.
CC -!- SIMILARITY: CONTAINS 2 HEMOPEXIN-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 SOMATOMEDIN-B TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

Search completed: January 29, 2002, 11:13:47
Job time: 823 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:12:14 ; Search time 285.36 seconds
(without alignments)
11.277 Million cell updates/sec

Title: US-09-763-397A-19

Perfect score: 131

Sequence: 1 EFTYMINFGRGQNYWEHPYQKS 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhmc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.rodent:*
- 12: sp.virus:*
- 13: sp.vertebrate:*
- 14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	131	100.0	526	5 Q9N9G1	Q9n9g1 plasmodium
2	131	100.0	526	5 Q9N9F8	Q9n9f8 plasmodium
3	131	100.0	526	5 Q9N9F7	Q9n9f7 plasmodium
4	131	100.0	526	5 Q9N9F5	Q9n9f5 plasmodium
5	131	100.0	526	5 Q9N9F3	Q9n9f3 plasmodium
6	131	100.0	526	5 Q9N9F2	Q9n9f2 plasmodium
7	131	100.0	526	5 Q9N9E4	Q9n9e4 plasmodium
8	131	100.0	526	5 Q9N9E3	Q9n9e3 plasmodium
9	131	100.0	526	5 Q9N9E0	Q9n9e0 plasmodium
10	131	100.0	592	5 Q25745	Q25745 plasmodium
11	131	100.0	592	5 Q25749	Q25749 plasmodium
12	131	100.0	604	5 Q25746	Q25746 plasmodium
13	131	100.0	605	5 Q9GVB7	Q9gvb7 plasmodium
14	131	100.0	620	5 Q00784	Q00784 plasmodium
15	131	100.0	621	5 Q9TY48	Q9ty48 plasmodium
16	131	100.0	622	5 Q9GZ05	Q9gz05 plasmodium
17	126	96.2	310	5 Q25750	Q25750 plasmodium
18	126	96.2	525	5 Q9N9F1	Q9n9f1 plasmodium
19	126	96.2	526	5 Q9N9G0	Q9n9g0 plasmodium

20	126	96.2	526	5 Q9N9F9	Q9n9f9 plasmodium
21	126	96.2	526	5 Q9N9F6	Q9n9f6 plasmodium
22	126	96.2	526	5 Q9N9E9	Q9n9e9 plasmodium
23	126	96.2	526	5 Q9N9E8	Q9n9e8 plasmodium
24	126	96.2	526	5 Q9N9E7	Q9n9e7 plasmodium
25	126	96.2	526	5 Q9N9E6	Q9n9e6 plasmodium
26	126	96.2	526	5 Q9N9E5	Q9n9e5 plasmodium
27	126	96.2	526	5 Q9N9E2	Q9n9e2 plasmodium
28	126	96.2	526	5 Q9N9E1	Q9n9e1 plasmodium
29	126	96.2	526	5 Q9N9D9	Q9n9d9 plasmodium
30	126	96.2	604	5 Q25747	Q25747 plasmodium
31	126	96.2	622	5 Q94661	Q94661 plasmodium
32	124	94.7	526	5 Q9N9F4	Q9n9f4 plasmodium
33	124	94.7	526	5 Q9N9F0	Q9n9f0 plasmodium
34	82	62.6	605	5 Q9NFW9	Q9nfw9 plasmodium
35	57	43.5	1455	4 Q9UPV0	Q9upv0 homo sapien
36	53.5	40.8	631	3 Q14353	Q14353 s blotin--p
37	53	40.5	737	2 Q9EYQ5	Q9eyq5 clostridium
38	52	39.7	506	10 Q9SSX5	Q9ssx5 antirrhinum
39	51	38.9	865	5 Q9N8I7	Q9n8i7 trypanosoma
40	51	38.9	2621	10 Q9LMZ3	Q9lmz3 arabidopsis
41	51	38.9	2658	10 Q9SGE4	Q9sge4 arabidopsis
42	49	37.4	218	10 Q49833	Q49833 nicotiana a
43	49	37.4	512	10 Q9SSX4	Q9ssx4 torenia hyb
44	48.5	37.0	115	10 Q9ZSG6	Q9zsg6 lycium ande
45	48	36.6	503	5 Q9VPI6	Q9vpi6 drosophila

ALIGNMENTS

RESULT 1
Q9N9G1 ID Q9N9G1 PRELIMINARY; PRT; 526 AA.
AC Q9N9G1:
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN 83/AMA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=44-2;
RX MEDLINE=20416492; PubMed=10960173;
RA Kocken C.H.M., Narum D.L., Massouboodji A., Ayivi B., Dubbeld M.A.,
van der Wel A., Conway D.J., Sanni A., Thomas A.W.;
RT "Molecular characterisation of Plasmodium reichenowi apical membrane
antigen-1 (AMA-1), comparison with P. falciparum AMA-1, and antibody-
mediated inhibition of red cell invasion";
RL Mol. Biochem. Parasitol. 109:147-156(2000).
DR EMBL; AJ271168; CAB97180.1; -;
DR InterPro; IPR003298; Apmem_Agl.
DR Pfam; PF02430; AMA-1; 1.
DR PRINTS; PR01361; MEROZOITESA.
FT NON_TER 1
FT NON_TER 526
SQ SEQUENCE 526 AA; 60680 MW; 2A05E3F84D1BEB9B CRC64;

Query Match 100.0%; Score 131; DB 5; Length 526;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFTYMINFGRGQNYWEHPYQKS 22
|||||
Db 6 EFTYMINFGRGQNYWEHPYQKS 27

RESULT 2
Q9N9F8

ID Q9N9F8 PRELIMINARY; PRT; 526 AA.
 AC Q9N9F8;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
 GN 83/AMA-1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C5-1.1;
 RX MEDLINE=20416492; PubMed=10960173;
 RA Kocken C.H.M., Narum D.L., Massougbodji A., Ayivi B., Dubbeld M.A.,
 van der Wel A., Conway D.J., Sanni A., Thomas A.W.;
 RT "Molecular characterisation of Plasmodium reichenowi apical membrane
 antigen-1 (AMA-1), comparison with P. falciparum AMA-1, and antibody-
 mediated inhibition of red cell invasion.";
 RL Mol. Biochem. Parasitol. 109:147-156(2000).
 DR EMBL; AJ271171; CAB97183.1; -;
 DR InterPro; IPR003298; Apmem_Ag1.
 DR Pfam; PF02430; AMA-1; 1.
 DR PRINTS; PR01361; MEROZOITESA.
 DR NON_TER 1 526
 FT NON_TER 1 526
 SQ SEQUENCE 526 AA; 60845 MW; FF0FABED213A7C4B CRC64;

Query Match 100.0%; Score 131; DB 5; Length 526;
 Best Local Similarity 100.0%; Pred. No. 1.2e-11;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFTYMINFGRGQNYWEHPYOKS 22
 |||||
 Db 6 EFTYMINFGRGQNYWEHPYOKS 27

RESULT 3
 ID Q9N9F7 PRELIMINARY; PRT; 526 AA.
 AC Q9N9F7;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
 GN 83/AMA-1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=38-12.1;
 RX MEDLINE=20416492; PubMed=10960173;
 RA Kocken C.H.M., Narum D.L., Massougbodji A., Ayivi B., Dubbeld M.A.,
 van der Wel A., Conway D.J., Sanni A., Thomas A.W.;
 RT "Molecular characterisation of Plasmodium reichenowi apical membrane
 antigen-1 (AMA-1), comparison with P. falciparum AMA-1, and antibody-
 mediated inhibition of red cell invasion.";
 RL Mol. Biochem. Parasitol. 109:147-156(2000).
 DR EMBL; AJ271172; CAB97184.1; -;
 DR InterPro; IPR003298; Apmem_Ag1.
 DR Pfam; PF02430; AMA-1; 1.
 DR PRINTS; PR01361; MEROZOITESA.
 DR NON_TER 1 526
 FT NON_TER 1 526
 SQ SEQUENCE 526 AA; 60955 MW; ALE2FDA9FFCA700F CRC64;

Query Match 100.0%; Score 131; DB 5; Length 526;
 Best Local Similarity 100.0%; Pred. No. 1.2e-11;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

~ ~ ~

QY 1 EFTYMINFGRGQNYWEHPYOKS 22
 |||||
 Db 6 EFTYMINFGRGQNYWEHPYOKS 27

RESULT 4
 ID Q9N9F5 PRELIMINARY; PRT; 526 AA.
 AC Q9N9F5;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
 GN 83/AMA-1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=33-6.1;
 RX MEDLINE=20416492; PubMed=10960173;
 RA Kocken C.H.M., Narum D.L., Massougbodji A., Ayivi B., Dubbeld M.A.,
 van der Wel A., Conway D.J., Sanni A., Thomas A.W.;
 RT "Molecular characterisation of Plasmodium reichenowi apical membrane
 antigen-1 (AMA-1), comparison with P. falciparum AMA-1, and antibody-
 mediated inhibition of red cell invasion.";
 RL Mol. Biochem. Parasitol. 109:147-156(2000).
 DR EMBL; AJ271174; CAB97186.1; -;
 DR InterPro; IPR003298; Apmem_Ag1.
 DR Pfam; PF02430; AMA-1; 1.
 DR PRINTS; PR01361; MEROZOITESA.
 DR NON_TER 1 526
 FT NON_TER 1 526
 SQ SEQUENCE 526 AA; 60933 MW; 12D4A685A55510F2 CRC64;

Query Match 100.0%; Score 131; DB 5; Length 526;
 Best Local Similarity 100.0%; Pred. No. 1.2e-11;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFTYMINFGRGQNYWEHPYOKS 22
 |||||
 Db 6 EFTYMINFGRGQNYWEHPYOKS 27

RESULT 5
 ID Q9N9F3 PRELIMINARY; PRT; 526 AA.
 AC Q9N9F3;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
 GN 83/AMA-1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=29-3.1;
 RX MEDLINE=20416492; PubMed=10960173;
 RA Kocken C.H.M., Narum D.L., Massougbodji A., Ayivi B., Dubbeld M.A.,
 van der Wel A., Conway D.J., Sanni A., Thomas A.W.;
 RT "Molecular characterisation of Plasmodium reichenowi apical membrane
 antigen-1 (AMA-1), comparison with P. falciparum AMA-1, and antibody-
 mediated inhibition of red cell invasion.";
 RL Mol. Biochem. Parasitol. 109:147-156(2000).
 DR EMBL; AJ271176; CAB97188.1; -;
 DR InterPro; IPR003298; Apmem_Ag1.
 DR Pfam; PF02430; AMA-1; 1.
 DR PRINTS; PR01361; MEROZOITESA.
 DR NON_TER 1 526
 FT NON_TER 1 526

DE APICAL MEMBRANE ANTIGEN 1 (FRAGMENT).
GN AMA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVO;
RA Kocken C.H.M., Dubbeld M.A., Herrera S., Thomas A.W.;
RT "Diversity of Apical Membrane Antigen 1 sequence in Plasmodium
falciparum FVO strain."; the EMBL/GenBank/DBJ databases.
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ277646; CAC05390.1; -.
DR InterPro: IPR003298; Apmem_Ag1.
DR Pfam: PF02430; AMA-1; 1.
DR PRINTS: PR01361; MEROZOITESA.
DR NON_TER 1
FT NON_TER 605
SQ SEQUENCE 605 AA; 69897 MW; F6D89F02F7690C19 CRC64;

Query Match 100.0%; Score 131; DB 5; Length 605;
Best Local Similarity 100.0%; Pred. No. 1.4e-11; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 EFTYMINFGRGQNYWEHPYQKS 22
|||||
Db 6 EFTYMINFGRGQNYWEHPYQKS 27

RESULT 14
C00784
ID O00784 PRELIMINARY; PRT; 620 AA.
AC O00784;
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN-1 (FRAGMENT).
GN AMA-1.
OS Plasmodium falciparum (isolate kf1916).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57269;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KF1916;
RA Eisen D., Billman-Jacobe H., Marshall V.M., Coppel R.L.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KF1916;
RA Eisen D.P., Marshall V.M., Billman-Jacobe H., Coppel R.L.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: U90276; AAB50405.1; -.
DR EMBL: AF061332; AAC15773.1; -.
DR InterPro: IPR003298; Apmem_Ag1.
DR Pfam: PF02430; AMA-1; 1.
FT NON_TER 620
SQ SEQUENCE 620 AA; 71721 MW; 9194F06A0D138677 CRC64;

Query Match 100.0%; Score 131; DB 5; Length 620;
Best Local Similarity 100.0%; Pred. No. 1.4e-11; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 EFTYMINFGRGQNYWEHPYQKS 22
|||||
Db 14 EFTYMINFGRGQNYWEHPYQKS 35

RESULT 15
Q9TY48
ID Q9TY48 PRELIMINARY; PRT; 621 AA.
AC Q9TY48;

DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE APICAL MEMBRANE ANTIGEN-1 (FRAGMENT).
GN AMA-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVO;
RA Aguilar J.C., Hoffman S.L.;
RT "The immunogenicity and protective efficacy of Plasmodium falciparum
AMA-1 DNA vaccines.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U84348; AAD03790.1; -.
DR InterPro: IPR003298; Apmem_Ag1.
DR Pfam: PF02430; AMA-1; 1.
FT NON_TER 621
SQ SEQUENCE 621 AA; 71921 MW; 6EA2A8C5142D8A7D CRC64;

Query Match 100.0%; Score 131; DB 5; Length 621;
Best Local Similarity 100.0%; Pred. No. 1.4e-11; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 EFTYMINFGRGQNYWEHPYQKS 22
|||||
Db 14 EFTYMINFGRGQNYWEHPYQKS 35

Search completed: January 29, 2002, 11:12:14
Job time: 770 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:21:50 : Search time 310.82 Seconds
(without alignments)
4.528 Million cell updates/sec

Title: US-09-763-397A-20

Perfect score: 96

Sequence: 1 NEREDETLTKYEYDVLK 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_1101.*

```

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	19	21	AA70296
2	96	100.0	350	21	AA70278
3	96	100.0	1435	16	AA70232
4	96	100.0	1435	18	AAW22477
5	96	100.0	1435	21	AA777900
6	96	100.0	1604	16	AA70105
7	96	100.0	1786	14	AA41043
8	59	61.5	21	18	AAW35466
9	54	56.2	21	18	AAW35507
10	48	48	587	17	AAW88656
11	48	50.0	679	20	AA701524

12	48	50.0	685	20	AA701525	TbpB sequence from
13	48	50.0	689	17	AA88648	Neisseria meningit
14	48	50.0	689	20	AA701523	TbpB sequence from
15	47	49.0	199	17	AA88661	N.meningitidis B28
16	45	46.9	245	22	AA13856	Peptide #290 encod
17	45	46.9	245	22	AA26263	Peptide #300 encod
18	45	46.9	245	22	AA01599	Peptide #281 encod
19	45	46.9	645	22	AA82945	S. epidermidis ope
20	45	46.9	1501	22	AA82878	S. epidermidis ope
21	44	45.8	517	18	AA30919	Tobacco calcium/ca
22	43	44.8	211	22	AA93906	C glutamicum prote
23	43	44.8	234	21	AA804280	Arabidopsis thalia
24	43	44.8	245	21	AA804279	Arabidopsis thalia
25	43	44.8	494	21	AA836673	Arabidopsis thalia
26	43	44.8	521	21	AA836672	Arabidopsis thalia
27	43	44.8	620	21	AA836671	Arabidopsis thalia
28	43	44.8	693	22	AA401118	Human polypeptide
29	43	44.8	1316	21	AA830505	A calcium-dependen
30	43	44.8	1337	21	AA830504	Plasmodium falcipa
31	43	44.8	2573	21	AA818234	S. epidermidis ope
32	42	43.8	265	22	AA82827	S. epidermidis ope
33	42	43.8	766	22	AA82742	D. immitis ankyrin
34	41	42.7	52	21	AA81628	D. immitis ankyrin
35	41	42.7	259	21	AA854134	Human pancreatic c
36	41	42.7	270	19	AAW70602	Ankyrin protein fr
37	41	42.7	270	19	AAW76770	D. immitis ankyrin
38	41	42.7	270	21	AA811583	D. immitis ankyrin
39	41	42.7	285	21	AA856678	Arabidopsis thalia
40	41	42.7	287	21	AA81618	D. immitis ankyrin
41	41	42.7	309	22	AA87763	Human T2R33 amino
42	41	42.7	312	19	AAW70601	Ankyrin protein fr
43	41	42.7	312	19	AAW76769	D. immitis ankyrin
44	41	42.7	312	21	AA811582	D. immitis ankyrin
45	41	42.7	371	21	AA856677	Arabidopsis thalia

ALIGNMENTS

```

RESULT 1
AA70296
ID  AA70296 standard; peptide; 19 AA.
XX
AC  AA70296;
XX
DT  06-JUN-2000 (first entry)
XX
DE  Plasmodium falciparum EBA-175 antigenic epitope, P346.
XX
KW  Recombinant protein; CDC/NIAIDVAC-1; multivalent; malaria; vaccine;
KW  T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
KW  circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
KW  liver stage antigen-1; USA-1; merozoite surface protein-1; MSP-1; MSP-2;
KW  apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
KW  EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
KW  Pf27; antiparasitic; prevention; anti-CDC/NIAIDVAC-1 antibody.
XX
OS  Plasmodium falciparum.
XX
PN  WO200011179-A1.
XX
PD  02-MAR-2000.
XX
PF  19-AUG-1999; 99WO-US18869.
XX
PR  21-AUG-1998; 98US-0097703.
XX
PA  (NATIM-) NAT INST IMMUNOLOGY.
PA  (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI  Lal AA, Shi YP, Hasnain SE;
XX  WPI; 2000-237654/20.
DR

```

```

XX Novel recombinant protein as vaccine for treating malarial infection
PT comprises antigenic peptides obtained from different stages of
PT plasmodium falciparum life cycle -
XX
XX Claim 2; Page 17; 52pp; English.
XX
XX The present sequence is the antigenic epitope P546, derived from
CC erythrocyte binding antigen-175 (EBA-175) of the asexual blood stage of
CC Plasmodium falciparum. It is used in the construction of recombinant
CC protein CDC/NIIMALVAC-1, which is a multivalent, multistage malarial
CC vaccine. The recombinant protein comprises, melittin signal peptide,
CC (His)6 tag, T-cell epitope from tetanus toxoid and 21 antigenic epitopes
CC from circumsporozoite protein (CSP), sporozoite surface protein-2
CC (SSP-2), liver stage antigen-1 (LSA-1), merozoite surface protein-1
CC (MSP-1), MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding
CC antigen-175 (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete
CC specific antigen, Pfg27. These epitopes were obtained at different stages
CC of the life cycle of P. falciparum. CDC/NIIMALVAC-1 vaccine has
CC antiparasitic activity and can be used for treatment and prevention of
CC malarial infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for
CC detecting P. falciparum in biological samples.
XX
XX Sequence 19 AA;
SQ
Query Match 100.0%; Score 96; DB 21; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEREDERTLTKEYEDIVLK 19
DB 1 neredertltkeyedivlk 19

RESULT 2
AAV70278
ID AAY70278 standard; Protein; 350 AA.
XX
XX AAY70278;
AC
XX
XX 06-JUN-2000 (first entry)
DT
XX
XX Recombinant vaccine CDC/NIIMALVAC-1.
DE
XX
XX Recombinant protein; CDC/NIIMALVAC-1; multivalent; malaria; vaccine;
KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
KW Pfg27; antiparasitic; prevention; anti-CDC/NIIMALVAC-1 antibody;
KW honey bee.
XX
XX Chimeric - Apis sp.
OS
XX Chimeric - Clostridium tetani.
OS
XX Chimeric - Plasmodium falciparum.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..22
FT /label= Melittin_signal_peptide
FT /note= "Derived from Honey bee"
FT 23..350
FT Protein
FT /label= Mature_CDC/NIIMALVAC-1
FT /note= "Recombinant multivalent malarial vaccine"
XX
XX WO200011179-A1.
XX
XX 02-MAR-2000.
XX
XX 19-AUG-1999; 99WO-US18869.
XX
XX 21-AUG-1998; 98US-0097703.
XX

```

```

XX (NAIM-) NAT INST IMMUNOLOGY.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Lal AA, Shi YP, Hasnain SE;
XX
XX WPI: 2000-237654/20.
DR N-PSDB; AA251336.
XX
XX Novel recombinant protein as vaccine for treating malarial infection
PT comprises antigenic peptides obtained from different stages of
PT plasmodium falciparum life cycle -
XX
XX Claim 3; Page 43-44; 52pp; English.
XX
XX The present sequence is that of recombinant protein CDC/NIIMALVAC-1,
CC which is a multivalent, multistage malarial vaccine. The recombinant
CC protein comprises, melittin signal peptide, (His)6 tag, T-cell epitope
CC from tetanus toxoid and 21 antigenic epitopes from circumsporozoite
CC protein (CSP), sporozoite surface protein-2 (SSP-2), liver stage
CC antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1), MSP-2, apical
CC membrane antigen-1 (AMA-1), erythrocyte binding antigen-175 (EBA-175),
CC rhoptry associated protein-1 (RAP-1) and gamete specific antigen, Pfg27.
CC These epitopes were obtained at different stages of the life cycle of
CC Plasmodium falciparum. CDC/NIIMALVAC-1 vaccine has antiparasitic
CC activity and can be used for treatment and prevention of malarial
CC infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for detecting
CC P. falciparum in biological samples.
XX
XX Sequence 350 AA;
SQ
Query Match 100.0%; Score 96; DB 21; Length 350;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEREDERTLTKEYEDIVLK 19
DB 223 neredertltkeyedivlk 241

RESULT 3
AAV70232
ID AAR70232 standard; Protein; 1435 AA.
XX
XX AAR70232;
AC
XX 22-SEP-1995 (first entry)
DT
XX P. falciparum SABP.
DE
XX
XX SABP; sialic acid binding protein; binding domain; merozoite;
KW malaria; therapy; vaccine.
KW
XX
XX Plasmodium falciparum.
OS
XX
XX Key Location/Qualifiers
FH Domain 1..616
FT /label= Binding_domain
FT
XX
XX WO9507353-A.
XX
XX 16-MAR-1995.
XX
XX 07-SEP-1994; 94WO-US10230.
XX
XX 10-SEP-1993; 93US-0119677.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Chitnis C, Miller LH, Peterson DS, Sim KL, Su X;
PI Wellens TE;
XX

```

DR WPI: 1995-123427/16.
 DR N-PSDB; AAQ83525.
 XX New erythrocyte binding domain polypeptide(s) - isolated from
 PT Plasmodium binding proteins, used in diagnosis, treatment and
 PT prevention of malaria
 XX
 PS Disclosure; Page 39-41; 81pp; English.
 XX Sequences from the SABP gene (given in AAQ83525) were PCR amplified,
 CC expressed on the surface of COS cells and tested for erythrocyte
 CC binding to identify the binding domain polypeptide. A pref. SABP
 CC binding domain comprises residues 1 to about 616 of the SABP protein
 CC (AAQ83525). Recombinant binding domain was expressed in E. coli,
 CC yeast, mammalian, insect, and in vaccinia virus and adenovirus-
 CC infected cells. It provides protection against P. falciparum.
 XX
 SQ Sequence 1435 AA;
 Query Match 100.0%; Score 96; DB 16; Length 1435;
 Best Local Similarity 100.0%; Pred. No. 5.5e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NEREDERTLTKEYEDIVLK 19
 ||||||||||||||||
 Db 1069 neredertltkeyedivlk 1087
 RESULT 4
 ID AAW22477 standard; Protein; 1435 AA.
 XX AC AAW22477;
 XX
 DT 11-SEP-1997 (first entry)
 XX
 DE Sialic acid binding protein.
 XX DBL gene family; SABP; sialic acid binding protein; merozoite; malaria;
 KW Duffy antigen binding protein; DABP; erythrocyte; var-1; var-2; var-3;
 KW var-7; vaccine; therapy; immune response; Plasmodium.
 XX Plasmodium falciparum.
 OS WO9640766-A2.
 PN 19-DEC-1996.
 XX
 PD 07-JUN-1996; 96WO-US09508.
 XX
 PF 07-JUN-1995; 95US-0487826.
 XX
 PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Chitnis C, Miller LH, Peterson DS, Sim KL, Su X;
 XX Wellem's TE;
 PI WPI: 1997-052231/05.
 XX N-PSDB; AAT72888.
 DR New malaria vaccines - contains cysteine-rich DBL family protein
 PT binding domains homologous domains of the Duffy and sialic acid
 PT binding proteins
 XX
 PS Example 1; Page 37-40; 96pp; English.
 XX This sequence represents the full length sialic acid binding protein
 CC (SABP). SABP and the Duffy antigen binding protein (DABP) are soluble
 CC proteins that appear in the culture supernatant after infected
 CC erythrocytes release merozoites. DABP and SABP mediate the binding of
 CC merozoites and schizonts to the erythrocyte surface. These proteins are
 CC necessary for erythrocyte invasion by the parasite. This sequence can be

CC used in the compositions of the invention. The compositions are for the
 CC treatment and prevention of malaria, and comprise either a nucleotide
 CC sequence or encoded polypeptide of the var-1, var-2, var-3 or var-7 genes
 CC of the DBL gene family, a family of genes having homology with conserved
 CC regions of DABP and SABP. The compositions are used for the treatment and
 CC prevention of malaria. They are also used in the preparation of vaccines
 CC for inducing a protective immune response in a mammal to Plasmodium
 CC merozoites (especially Plasmodium falciparum or Plasmodium vivax).
 XX
 SQ Sequence 1435 AA;
 Query Match 100.0%; Score 96; DB 18; Length 1435;
 Best Local Similarity 100.0%; Pred. No. 5.5e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NEREDERTLTKEYEDIVLK 19
 ||||||||||||||||
 Db 1069 neredertltkeyedivlk 1087
 RESULT 5
 ID AAY77900 standard; Protein; 1435 AA.
 XX AC AAY77900;
 XX
 DT 13-JUN-2000 (first entry)
 XX
 DE P. falciparum SABP binding domain polypeptide.
 XX DBL gene; Duffy-binding like gene; ebl-1; Duffy Antigen Binding Protein;
 KW DABP; Sialic Acid Binding Protein; SABP; malaria; vaccine; immunisation;
 KW protozoacide.
 XX Plasmodium falciparum.
 OS US5993827-A.
 PN 30-NOV-1999.
 PD 07-JUN-1995; 95US-0487826.
 XX
 PF 10-SEP-1993; 93US-0119677.
 XX
 PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Sim KL, Chitnis C, Peterson DS, Su X, Wellem's TE, Miller LH;
 XX WPI: 2000-194198/17.
 DR N-PSDB; AA298282.
 XX Isolated protein binding domains from Plasmodium vivax and Plasmodium
 PT falciparum erythrocyte binding proteins useful for vaccinating against
 PT malaria -
 XX Example 1; Columns 45-52; 93pp; English.
 PS The invention relates to ebl-1 polypeptides that are encoded by the DBL
 CC (Duffy-binding like) gene family. The ebl-1 proteins are substantially
 CC identical to the Duffy Antigen Binding Protein (DABP) and Sialic Acid
 CC Binding Protein (SABP), which are soluble proteins that appear in the
 CC culture supernatant after erythrocytes infected with malaria release
 CC merozoites. Immunochemical studies indicate that DABP and SABP are the
 CC respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy
 CC and sialic acid receptors on erythrocytes. The ebl-1 polypeptides may be
 CC used to vaccinate against malaria, especially caused by P. falciparum.
 CC Immunization with the polypeptide provides effective protection against
 CC malaria. The present sequence represents the SABP binding domain
 CC polypeptide.
 XX Sequence 1435 AA;

```
QY      1 NEREDERTLTKEYEDIVLK 19
        |||
Db     1420 neredertltkeyedivlk 14
```

```
RESULT 8
AAW35466
ID AAW35466 standard; peptide: 21 AA.
XX
AC AAW35466;
XX
DT 22-APR-1998 (first entry)
XX
DE Immunogenic agent peptide EBA 175-peptide.
XX
KW T-cell stimulatory peptide; immunogen; non-dendritic; carrier; tumour;
KW scaffold; inhibition; metastasis; wound healing; solid phase.
XX
OS Unidentified.
XX
PN WO9738011-A1.
XX
PD 16-OCT-1997.
XX
PF 03-APR-1997; 97WO-DE00146.
XX
PR 03-APR-1996; 96DK-0000398.
XX
PA (PEPR-) PEPRESEARCH AS.
XX
PI Heegaard PMH, Jakobsen PH;
XX
DR WPI; 1997-512645/47.
XX
PT Non-dendritic peptide carrier linked to a solid phase - useful as a
PT diagnostic agent and as a scaffold for production of chemical
PT derivatives
XX
PS Claim 105; Page 209; 262pp; English.
XX
CC A non-dendritic peptide carrier (A) has been developed which is coupled
CC through a linker to a solid phase, forming a complex of (A)-solid phase.
CC Where (A) comprises 10-50 amino acids capable of forming a secondary
CC structure in a benign buffer after liberation from the solid phase, and
CC further the (A)-solid phase complex comprises an immunogenic substance
CC and/or an immune mediator coupled on (A). The present sequence
CC represents a peptide used in an example from the present invention. An
CC (A)-solid phase complex can be used as a scaffold for the production of
CC chemical derivatives, characterised by covalently attaching molecules at
CC attachment points. Alternatively (A) is used as a scaffold-peptide for
CC the incorporation into an immunostimulating complex (Iscom) resulting an
CC (A)-Iscom complex which is used for the chemical coupling of antigenic
CC substances in an aqueous solution by conjugation. (A) derivatised with
CC one or more peptides having fibronectin-, laminin- or vitronectin-like
CC binding activities can be used for the promotion of cell-attachment to
CC plastic surfaces, in particular to inhibit tumour growth and metastasis,
CC and for promotion of wound healing. Also a derivatised (A) can be used
CC for the selection of specifically-binding aptamers or as a diagnostic
CC agent. Such diagnostic-(A) molecules could be used to detect molecules
CC derived from or indicative of pregnancy or of a disease, such as an
CC infectious, autoimmune or cancerous disease.
XX
SQ Sequence 21 AA;

Query Match 61.5%; Score 59; DB 18; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 TLTKEYEDIVLK 19
DB 1 titkeyedivlk 12

RESULT 9
AAW35507
ID AAW35507 standard; peptide: 21 AA.
```

```
XX
AC AAW35507;
XX
DT 22-APR-1998 (first entry)
XX
DE Erythrocyte bonding antigen-175 peptide.
XX
KW T-cell stimulatory peptide; immunogen; non-dendritic; carrier; tumour;
KW scaffold; inhibition; metastasis; wound healing; solid phase.
XX
OS Unidentified.
XX
PN WO9738011-A1.
XX
PD 16-OCT-1997.
XX
PF 03-APR-1997; 97WO-DE00146.
XX
PR 03-APR-1996; 96DK-0000398.
XX
PA (PEPR-) PEPRESEARCH AS.
XX
PI Heegaard PMH, Jakobsen PH;
XX
DR WPI; 1997-512645/47.
XX
PT Non-dendritic peptide carrier linked to a solid phase - useful as a
PT diagnostic agent and as a scaffold for production of chemical
PT derivatives
XX
PS Example 20; Page 124; 262pp; English.
XX
CC A non-dendritic peptide carrier (A) has been developed which is coupled
CC through a linker to a solid phase, forming a complex of (A)-solid phase.
CC Where (A) comprises 10-50 amino acids capable of forming a secondary
CC structure in a benign buffer after liberation from the solid phase, and
CC further the (A)-solid phase complex comprises an immunogenic substance
CC and/or an immune mediator coupled on (A). The present sequence
CC represents a peptide used in an example from the present invention. An
CC (A)-solid phase complex can be used as a scaffold for the production of
CC chemical derivatives, characterised by covalently attaching molecules at
CC attachment points. Alternatively (A) is used as a scaffold-peptide for
CC the incorporation into an immunostimulating complex (Iscom) resulting an
CC (A)-Iscom complex which is used for the chemical coupling of antigenic
CC substances in an aqueous solution by conjugation. (A) derivatised with
CC one or more peptides having fibronectin-, laminin- or vitronectin-like
CC binding activities can be used for the promotion of cell-attachment to
CC plastic surfaces, in particular to inhibit tumour growth and metastasis,
CC and for promotion of wound healing. Also a derivatised (A) can be used
CC for the selection of specifically-binding aptamers or as a diagnostic
CC agent. Such diagnostic-(A) molecules could be used to detect molecules
CC derived from or indicative of pregnancy or of a disease, such as an
CC infectious, autoimmune or cancerous disease.
XX
SQ Sequence 21 AA;

Query Match 56.2%; Score 54; DB 18; Length 21;
Best Local Similarity 91.7%; Pred. No. 0.089;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 TLTKEYEDIVLK 19
DB 1 tctkeyedivlk 12

RESULT 10
AAW88656
ID AAW88656 standard; Protein; 587 AA.
XX
AC AAW88656;
XX
DT 03-SEP-1996 (first entry)
```

Matches	9;	Conservative	4;	Mismatches	5;	Indels	0;	Gap
Qy	2	EREDERTLTKKEYEDIVLK 19						
		: : : :						
Db	424	ertdeikeipkeqqdivyr 441						
RESULT 11								
AAV01524								
ID	AAV01524	standard; Protein; 679 AA.						
XX	XX							
AC	AAV01524;							
DT	14-JUN-1999	(first entry)						
DE	TtpB	sequence from Neisseria meningitidis strain M982.						
KW	Low molecular weight subunit; TtpB;	human transferrin receptor; hTR;						
KW	Neisseria meningitidis; vaccine;	meningococcal infection; meningitis						
OS	Neisseria meningitidis.							
XX	XX							
PN	WO9907741-A1.							
PD	18-FEB-1999.							
XX	XX							
PF	03-AUG-1998;	98WO-FR01730.						
XX	XX							
PR	07-AUG-1997;	97FR-0010301.						
XX	XX							
PA	(INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.							
XX	XX							
PI	Rokbi B, Quentin-Millet M;							
XX	XX							
DR	WPI: 1999-190036/16.							
DR	N-PSDB; AAX26576.							
XX	XX							
PT	Vaccine containing small subunit of human transferrin receptor from							
PT	Neisseria meningitidis - for treatment and prevention of meningitis							
XX	XX							
PS	Claim 11; Page 38-40; 73pp; French.							
XX	XX							
CC	The present sequence represents the low molecular weight subunit							
CC	(TtpB) of the human transferrin receptor (hTR) from a specific							
CC	strain of Neisseria meningitidis that contains TtpB-encoding DNA.							
CC	Compositions containing TtpB are used as vaccines for treatment or							
CC	prevention of meningococcal infections, particularly meningitis.							
XX	XX							
SQ	Sequence	679 AA;						
Query Match	50.0%;	Score 48;	DB 20;	Length 679;				
Best Local Similarity	50.0%;	Pred. No. 41;						
Matches	9;	Conservative	4;	Mismatches	5;	Indels	0;	Gaps
Qy	2	EREDERTLTKKEYEDIVLK 19						
		: : : :						
Db	516	ertdeikeipkeqqdivyr 533						
RESULT 12								
AAV01525								
ID	AAV01525	standard; Protein; 685 AA.						
XX	XX							
AC	AAV01525;							
DT	14-JUN-1999	(first entry)						
DE	TtpB	sequence from Neisseria meningitidis strain 8680.						
XX	XX							
KW	Low molecular weight subunit; TtpB;	human transferrin receptor; hTR;						
KW	Neisseria meningitidis; vaccine;	meningococcal infection; meningitis						
XX	XX							

OS Neisseria meningitidis.

XX W09907741-A1.

XX 18-FEB-1999.

XX 03-AUG-1998; 98WO-FR01730.

XX 07-AUG-1997; 97FR-0010301.

XX (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

XX Rokbi B, Quentin-Millet M;

XX WPI; 1999-190036/16.

XX N-PSDB; AAX26577.

XX Vaccine containing small subunit of human transferrin receptor from

PT Neisseria meningitidis - for treatment and prevention of meningitis

XX Claim 10; Page 43-46; 73pp; French.

XX The present sequence encodes the low molecular weight subunit

CC (TbpB) of the human transferrin receptor (hTR) from a specific

CC strain of Neisseria meningitidis that contains TbpB-encoding DNA.

CC Compositions containing TbpB are used as vaccines for treatment or

CC prevention of meningococcal infections, particularly meningitis.

XX SQ Sequence 685 AA;

Query Match 50.0%; Score 48; DB 20; Length 685;

Best Local Similarity 50.0%; Pred. No. 42;

Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 EREDERTLTKEYEDIVLK 19

Db 522 ertdekeipkeqgdivr 539

II II : II : III :

RESULT 13

AAR88648

ID AAR88648 standard; Protein: 689 AA.

XX AAR88648;

XX 02-SEP-1996 (first entry)

XX Neisseria meningitidis B283 transferrin receptor Tbp2 subunit.

XX Transferrin receptor; Tbp2 subunit; deletion mutant; vaccine;

XX passive immunisation; immunotherapy; IM2169; IM2394.

XX Neisseria meningitidis (strain B283).

XX Key Location/Qualifiers

XX Protein 1..689

XX Peptide /label= Tbp2_subunit

XX Domain 1..20

XX Domain 21..342

XX Domain /label= Domain_1

XX Domain /note= "N-terminal domain"

XX Domain 343..541

XX Domain /label= Domain_2

XX Domain /note= "hinge domain"

XX Domain 542..689

XX Domain /label= Domain_3

XX Domain /note= "C-terminal domain"

XX W09533049-A2.

XX 07-DEC-1995.

XX 30-MAY-1995; 95WO-FR00701.

XX 31-MAY-1994; 94FR-0006594.

XX (INMR) PASTEUR MERIEUX SERUMS & VACCINS.

XX (TRGE) TRANSGENE SA.

XX Jacobs E, Legrain M, Mazarin V, Lissolo L;

XX Millet MBJ;

XX WPI; 1996-030562/03.

XX N-PSDB; AAT11243.

XX Polypeptide(s) for vaccination against Neisseria meningitidis group

PT B - comprising deletion mutants of transferrin receptor Tbp2

PT subunit

XX Disclosure; Page 76-81; 114pp; French.

XX The present sequence is that of the N.meningitidis strain B283

CC transferrin Tbp2 subunit. The Tbp2 polypeptide has three

CC domains (see features table); deletion mutants in which at least

CC one of the domains is partially or totally deleted are claimed,

CC provided that the first and second domains are not simultaneously

CC partially or totally deleted. The positions of the 3 domains in

CC B283 are defined by alignment with the IM2169 sequence. The

CC deletion mutant polypeptides of the invention can generate an

CC immune response against N.meningitidis.

XX SQ Sequence 689 AA;

Query Match 50.0%; Score 48; DB 17; Length 689;

Best Local Similarity 50.0%; Pred. No. 42;

Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 EREDERTLTKEYEDIVLK 19

Db 526 ertdekeipkeqgdivr 543

II II : II : III :

RESULT 14

AAY01523

ID AAY01523 standard; Protein: 689 AA.

XX AAY01523;

XX 14-JUN-1999 (first entry)

XX TbpB sequence from Neisseria meningitidis strain B283.

XX Low molecular weight subunit; TbpB; human transferrin receptor; hTR;

XX Neisseria meningitidis; vaccine; meningococcal infection; meningitis.

XX Neisseria meningitidis.

XX W09907741-A1.

XX 18-FEB-1999.

XX 03-AUG-1998; 98WO-FR01730.

XX 07-AUG-1997; 97FR-0010301.

XX (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

XX Rokbi B, Quentin-Millet M;

XX WPI; 1999-190036/16.

XX N-PSDB; AAX26575.

XX Vaccine containing small subunit of human transferrin receptor from

PT

PT Neisseria meningitidis - for treatment and prevention of meningitis
 XX Claim 4; Page 32-34; 73pp; French.
 XX The present sequence represents the low molecular weight subunit
 CC (TbpB) of the human transferrin receptor (hTR) from a specific
 CC strain of Neisseria meningitidis that contains TbpB-encoding DNA.
 CC Compositions containing TbpB are used as vaccines for treatment or
 CC prevention of meningococcal infections, particularly meningitis.
 XX
 SQ Sequence 689 AA;

Query Match 50.0%; Score 48; DB 20; Length 689;
 Best Local Similarity 50.0%; Pred. No. 42;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 EREDERTLTKEYEDIVLK 19
 || || : || : || :
 Db 526 ertdekeipkeggdivyr 543

RESULT 15

AAR88661
 ID AAR88661 standard; Protein; 199 AA.

XX
 AC AAR88661;

XX
 DT 03-SEP-1996 (first entry)

XX
 DE N.meningitidis B283 Tbp2 domain 1 and domain 3 deletion mutant.

XX
 KW Transferrin receptor; Tbp2 subunit; deletion mutant; vaccine;

XX
 KW passive immunisation; immunotherapy; domain 2...

XX
 OS Neisseria meningitidis (strain B283).

OS Synthetic.

XX
 FH Key Location/Qualifiers

FT Domain 1..199

FT /label= Domain_2

FT /note= "complete hinge domain; the entire domain 1
 and domain 3 sequences have been deleted"

XX
 PN WO9533049-A2.

XX
 PD 07-DEC-1995.

XX
 PF 30-MAY-1995; 95WO-FR00701.

XX
 PR 31-MAY-1994; 94FR-0006594.

XX
 PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS.

PA (TRGE) TRANSGENE SA.

XX
 PI Jacobs E, Legrain M, Mazarin V, Lissolo L;

PI Millet MBJ;

XX
 DR WPI; 1996-030562/03.

XX
 PT Polypeptide(s) for vaccination against Neisseria meningitidis group
 PT B - comprising deletion mutants of transferrin receptor Tbp2
 subunit

XX
 PS Claim 1; Page -; 114pp; French.

XX
 CC Deletion mutants derived from the N.meningitidis strain IM2169
 CC transferrin Tbp2 subunit are claimed in which at least one of the
 CC three domains is partially or totally deleted, provided that the
 CC first and second domains are not simultaneously partially or totally
 CC deleted. Derivatives of IM2169-related strains are also claimed and
 CC for these polypeptides the positions of the 3 domains are defined by
 CC alignment with the IM2169 sequence. The deletion mutant polypeptides

CC of the invention can generate an immune response against
 CC N.meningitidis. The present sequence is that of a preferred mutant
 CC in which domains 1 and 3 of the IM2169-related strain B283 have been
 CC deleted in their entirety, i.e. it is the sequence of domain 2 of
 CC the B283 Tbp2 subunit.

XX Sequence 199 AA;

Query Match 49.0%; Score 47; DB 17; Length 199;
 Best Local Similarity 56.2%; Pred. No. 14;
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 EREDERTLTKEYEDIV 17

|| || : || : || : ||

Db 184 ertdekeipkeggdiv 199

Search completed: January 29, 2002, 10:21:51
 Job time: 427 sec

THIS PAGE BLANK (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:24:10 ; Search time 133.18 Seconds
(without alignments)
3.210 Million cell updates/sec

Title: US-09-763-397A-20

Perfect score: 96

Sequence: 1 NEREDRTLTKKEYDIVLK 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCRUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	1435	2	US-08-568-459A-4
2	96	100.0	1435	2	US-08-487-826B-4
3	41	42.7	52	3	US-09-065-474-170
4	41	42.7	270	2	US-09-031-485-7
5	41	42.7	270	2	US-08-847-429A-7
6	41	42.7	270	3	US-09-065-474-7
7	41	42.7	287	3	US-09-065-474-145
8	41	42.7	312	2	US-09-031-485-2
9	41	42.7	312	2	US-08-847-429A-2
10	41	42.7	312	3	US-09-065-474-2
11	41	42.7	1745	2	US-09-031-485-33
12	41	42.7	1745	2	US-08-847-429A-33
13	41	42.7	1745	3	US-09-065-474-33
14	40	41.7	25	1	US-07-718-274A-18
15	40	41.7	25	1	US-08-149-106-18
16	40	41.7	25	1	US-08-298-021-18
17	40	41.7	25	1	US-08-278-729A-9
18	40	41.7	101	1	US-08-155-343A-9
19	40	41.7	101	1	US-08-406-672-9
20	40	41.7	101	1	US-08-335-583C-42
21	40	41.7	101	1	US-08-643-563A-9
22	40	41.7	101	1	US-08-643-763A-9
23	40	41.7	101	1	US-08-462-623-9
24	40	41.7	101	1	US-08-451-953A-9
25	40	41.7	101	2	US-08-445-468A-9
26	40	41.7	101	2	US-08-461-397A-9
27	40	41.7	101	2	US-08-912-088-9

28 40 41.7 101 3 US-08-278-730A-9 Sequence 9, Appli
29 40 41.7 101 3 US-08-478-097A-10 Sequence 10, Appli
30 40 41.7 101 3 US-08-445-467-9 Sequence 9, Appli
31 40 41.7 101 3 US-08-480-515A-9 Sequence 9, Appli
32 40 41.7 101 3 US-08-289-222E-23 Sequence 23, Appli
33 40 41.7 101 4 US-08-414-033A-9 Sequence 9, Appli
34 40 41.7 101 4 US-09-054-526B-23 Sequence 23, Appli
35 40 41.7 101 4 US-08-271-556A-7 Sequence 7, Appli
36 40 41.7 101 4 US-08-931-858E-156 Sequence 156, App
37 40 41.7 101 4 US-08-981-739-156 Sequence 156, App
38 40 41.7 101 4 US-08-440-894A-9 Sequence 9, Appli
39 40 41.7 101 5 PCT-US92-01968-9 Sequence 9, Appli
40 40 41.7 101 5 PCT-US93-07190-9 Sequence 9, Appli
41 40 41.7 101 5 PCT-US93-07231-9 Sequence 9, Appli
42 40 41.7 101 5 PCT-US93-08742-9 Sequence 9, Appli
43 40 41.7 101 5 PCT-US93-08808-9 Sequence 9, Appli
44 40 41.7 101 5 PCT-US93-08885-9 Sequence 9, Appli
45 40 41.7 102 2 US-08-288-508C-14 Sequence 14, Appli

ALIGNMENTS

RESULT 1

US-08-568-459A-4

; Sequence 4, Application US/08568459A

; Patent No. 5849306

; GENERAL INFORMATION:

; APPLICANT: Sim, Kim L.

; APPLICANT: Chitnis, Chetan

; APPLICANT: Miller, Louis H.

; APPLICANT: Peterson, David S.

; APPLICANT: Su, Xin-zhaun

; APPLICANT: Wellens, Thomas E.

; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe Martens Olson & Bear

; STREET: 620 Newport Center Drive 16th Floor

; CITY: Newport Beach

; STATE: California

; COUNTRY: US

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/568,459A

; FILING DATE: 07-DEC-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Israelsen, Ned

; REGISTRATION NUMBER: 29,655

; REFERENCE/DOCKET NUMBER: NIH121.001CP1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 235-8550

; TELEFAX: (619) 235-0176

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1435 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; ORGANISM: Plasmodium falciparum

; US-08-568-459A-4

```
Query Match      100.0%; Score 96; DB 2; Length 1435;
Best Local Similarity 100.0%; Pred. No. 7.1e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEREDERTLKKEYEDIVLK 19
   |||||
Db 1069 NEREDERTLKKEYEDIVLK 1087

RESULT 2
US-08-487-826B-4
; Sequence 4, Application US/08487826B
; Patent No. 593827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1435 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; US-08-487-826B-4

Query Match      100.0%; Score 96; DB 2; Length 1435;
Best Local Similarity 100.0%; Pred. No. 7.1e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEREDERTLKKEYEDIVLK 19
   |||||
Db 1069 NEREDERTLKKEYEDIVLK 1087

RESULT 3
US-09-065-474-170
; Sequence 170, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
```

```
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-09-065-474-170

Query Match      42.7%; Score 41; DB 3; Length 52;
Best Local Similarity 61.5%; Pred. No. 6.5;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 DERTLTKEYEDIV 17
   :|||: |||
Db 4 EERTITVEDDV 16

RESULT 4
US-09-031-485-7
; Sequence 7, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
```

; APPLICATION NUMBER: US/09/031.485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA: US/08/847.429
; APPLICATION NUMBER: US/08/847.429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-485-7

Query Match 42.7%; Score 41; DB 2; Length 270;
Best Local Similarity 61.5%; Pred. No. 40;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 DERTLTKEYEDIV 17
:||||:| ||| |
Db 132 EERTITTYEDDV 144

RESULT 5
US-08-847-429A-7
; Sequence 7, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847.429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-847-429A-7

Query Match 42.7%; Score 41; DB 2; Length 270;
Best Local Similarity 61.5%; Pred. No. 40;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 DERTLTKEYEDIV 17
:||||:| ||| |
Db 132 EERTITTYEDDV 144

RESULT 6
US-09-065-474-7
; Sequence 7, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065.474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-c1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-065-474-7

Query Match 42.7%; Score 41; DB 3; Length 270;
Best Local Similarity 61.5%; Pred. No. 40;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 DERTLTKEYEDIV 17
:||||:| ||| |
Db 132 EERTITTYEDDV 144

RESULT 7
US-09-065-474-145
; Sequence 145, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND

```

; TITLE OF INVENTION:  USES THEREOF
; NUMBER OF SEQUENCES:  171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:  Carol Talkington Verser, Ph.D.
; ADDRESSEE:  Heskia Corporation
; STREET:  1825 Sharp Point Drive
; CITY:  Fort Collins
; STATE:  Colorado
; COUNTRY:  USA
; ZIP:  80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Floppy disk
; COMPUTER:  IBM PC compatible
; OPERATING SYSTEM:  Windows 95
; SOFTWARE:  WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  US/09/065,474
; FILING DATE:  24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME:  Verser, Carol Talkington
; REGISTRATION NUMBER:  37,459
; REFERENCE/DOCKET NUMBER:  HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:  970/493-7272
; TELEFAX:  970/484-9505
; INFORMATION FOR SEQ ID NO:  145:
; SEQUENCE CHARACTERISTICS:
; LENGTH:  287 amino acids
; TYPE:  amino acid
; TOPOLOGY:  linear
; MOLECULE TYPE:  Protein
; US-09-065-474-145

Query Match          42.7%;  Score 41;  DB 3;  Length 287;
Best Local Similarity 61.5%;  Pred. No. 43;
Matches 8;  Conservative 2;  Mismatches 3;  Indels 0;  Gaps 0;

QY      5 DERTLTKEYEDIV 17
       :|||:| |||
Db      149 EERTITTYEDDV 161

RESULT      8
US-09-031-485-2
; Sequence 2, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT:  Tang, Liang
; APPLICANT:  Blehm, E. Scot
; TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION:  USES THEREOF
; NUMBER OF SEQUENCES:  85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:  Carol Talkington Verser, Ph.D.
; ADDRESSEE:  Heskia Corporation
; STREET:  1825 Sharp Point Drive
; CITY:  Fort Collins
; STATE:  Colorado
; COUNTRY:  USA
; ZIP:  80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Floppy disk
; COMPUTER:  IBM PC compatible
; OPERATING SYSTEM:  Windows 95
; SOFTWARE:  WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  US/09/031,485
; FILING DATE:
; CLASSIFICATION:  530
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER:  US/08/847,429
; FILING DATE:  24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME:  Verser, Carol Talkington
; REGISTRATION NUMBER:  37,459
; REFERENCE/DOCKET NUMBER:  HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:  970/493-7272
; TELEFAX:  970/484-9505
; INFORMATION FOR SEQ ID NO:  2:
; SEQUENCE CHARACTERISTICS:
; LENGTH:  312 amino acids
; TYPE:  amino acid
; TOPOLOGY:  linear
; MOLECULE TYPE:  protein
; US-09-031-485-2

Query Match          42.7%;  Score 41;  DB 2;  Length 312;
Best Local Similarity 61.5%;  Pred. No. 48;
Matches 8;  Conservative 2;  Mismatches 3;  Indels 0;  Gaps 0;

QY      5 DERTLTKEYEDIV 17
       :|||:| |||
Db      297 EERTITTYEDDV 309

RESULT      9
US-08-847-429A-2
; Sequence 2, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT:  Tang, Liang
; APPLICANT:  Blehm, E. Scot
; TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION:  USES THEREOF
; NUMBER OF SEQUENCES:  85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:  Carol Talkington Verser, Ph.D.
; ADDRESSEE:  Heskia Corporation
; STREET:  1825 Sharp Point Drive
; CITY:  Fort Collins
; STATE:  Colorado
; COUNTRY:  USA
; ZIP:  80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Floppy disk
; COMPUTER:  IBM PC compatible
; OPERATING SYSTEM:  Windows 95
; SOFTWARE:  WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  US/08/847,429A
; FILING DATE:  24-APR-1997
; CLASSIFICATION:  435
; ATTORNEY/AGENT INFORMATION:
; NAME:  Verser, Carol Talkington
; REGISTRATION NUMBER:  37,459
; REFERENCE/DOCKET NUMBER:  HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:  970/493-7272
; TELEFAX:  970/484-9505
; INFORMATION FOR SEQ ID NO:  2:
; SEQUENCE CHARACTERISTICS:
; LENGTH:  312 amino acids
; TYPE:  amino acid
; TOPOLOGY:  linear
; MOLECULE TYPE:  protein
; US-08-847-429A-2

Query Match          42.7%;  Score 41;  DB 2;  Length 312;
Best Local Similarity 61.5%;  Pred. No. 48;

```


Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 5 DERTLTKEYEDIV 17
:||||:| ||| |
Db 297 EERTITTYEDDV 309

RESULT 10

US-09-065-474-2
; Sequence 2, Application US/09065474
; Patent No. 5063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065.474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 312 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-065-474-2

Query Match 42.7%; Score 41; DB 3; Length 312;
Best Local Similarity 61.5%; Pred. No. 48;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 5 DERTLTKEYEDIV 17
:||||:| ||| |
Db 297 EERTITTYEDDV 309

RESULT 11

US-09-031-485-33
; Sequence 33, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.

; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031.485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847.429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-485-33

Query Match 42.7%; Score 41; DB 2; Length 1745;
Best Local Similarity 61.5%; Pred. No. 3.2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 5 DERTLTKEYEDIV 17
:||||:| ||| |
Db 1607 EERTITTYEDDV 1619

RESULT 12

US-08-847-429A-33
; Sequence 33, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847.429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington

```
Query Match      42.7%; Score 41; DB 3; Length 1745;
Best Local Similarity 61.5%; Pred. NO. 3.2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels
```

```

RESULT 15
US-08-149-106-18
Sequence 18, Application US/08149106
Patent No. 5411941
GENERAL INFORMATION:
APPLICANT: Grinna, Lynn
APPLICANT: Parsons, Thomas F.
APPLICANT: Theofan, Georgia
TITLE OF INVENTION: Osteogenic Factor
NUMBER OF SEQUENCES: 63

```

Search completed: January 29, 2002, 10:24:11
Job time: 517 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:26:46 ; Search time 144.96 seconds
(without alignments)
9.984 Million cell updates/sec

Title: US-09-763-397A-20

Perfect score: 96

Sequence: 1 NEREDERTLTKEYEDIVLK 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	96	100.0	1435	2	A37793	erythrocyte-bindin
2	50	52.1	368	2	H84676	hypothetical prote
3	49	51.0	442	2	A36469	bam protein - frui
4	49	51.0	598	2	T38241	probable U3 small
5	48	50.0	217	2	S61546	transferrin bindin
6	48	50.0	689	2	S70661	transferrin-bindin
7	46	47.9	213	1	Q0V215	D9R protein - vacc
8	46	47.9	213	1	F36847	F9R protein - vari
9	46	47.9	213	2	H42515	D9R protein - vacc
10	46	47.9	213	2	A72163	D9R protein - vari
11	46	47.9	213	2	T28537	hypothetical prote
12	46	47.9	213	2	T37382	25K mutT-like prot
13	46	47.9	384	2	D84509	hypothetical prote
14	46	47.9	860	2	T37768	probable vacuolar
15	46	47.9	8243	2	T31307	type I fatty acid
16	45	46.9	256	2	S39969	tubulin alpha chai
17	45	46.9	513	2	S39691	UTP-hexose-1-phos
18	45	46.9	1282	2	T30577	DNA topoisomerase
19	44.5	46.4	361	2	T31242	hypothetical prote
20	44	45.8	333	1	A38094	D-lactate dehydrog
21	44	45.8	978	2	H81311	transcription-repa
22	44	45.8	985	2	S59330	Na+/H+-exchanging
23	43.5	45.3	731	2	S29870	DNA-binding protei
24	43	44.8	319	2	D81069	glutathione synthet
25	43	44.8	342	2	H81404	S-adenosylmethioni
26	43	44.8	943	2	S59317	DIP2 protein - yea
27	43	44.8	1829	2	T24583	hypothetical prote
28	43	44.8	2573	2	D71614	hypothetical prote
29	42.5	44.3	604	2	T15514	hypothetical prote

ALIGNMENTS

RESULT 1

A37793

erythrocyte-binding antigen 175 - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 09-Jun-2000

C:Accession: A37793; S11561

R:Sim, B.K.L.; Orlandi, P.A.; Haynes, J.D.; Klotz, F.W.; Carter, J.M.; Camus, D.; Zeg

J. Cell Biol. 111, 1877-1884, 1990

A:Title: Primary structure of the 175K Plasmodium falciparum erythrocyte binding anti

A:Reference number: A37793; MUID:91035602

A:Accession: A37793

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1435 <SIW>

A:Cross-references: GB:X52524; NID:g9884; PID:g9885

R:Sim, B.K.L.

Mol. Biochem. Parasitol. 41, 293-295, 1990

A:Title: Sequence conservation of a functional domain of erythrocyte binding antigen

A:Reference number: S11561; MUID:90377299

A:Accession: S11561

A:Molecule type: DNA

A:Residues: 995-1158 <SIW>

A:Cross-references: EMBL:X52524

Query Match 100.0%; Score 96; DB 2; Length 1435;

Best Local Similarity 100.0%; Pred. No. 2e-06;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEREDERTLTKEYEDIVLK 19

DB 1069 NEREDERTLTKEYEDIVLK 1087

RESULT 2

H84676

hypothetical protein At2g27780 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Mar-2001

C:Accession: H84676

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487

A:Accession: H84676

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-368 <STO>

A:Cross-references: GB:AE002093; NID:g3860254; PIDN:AAC73022.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g27780
A:Map position: 2
C:Superfamily: Arabidopsis thaliana hypothetical protein At2g32820

Query Match 52.1%; Score 50; DB 2; Length 368;
Best Local Similarity 56.2%; Pred. No. 4.7;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 NEREDERTLTKEYEDV 16
|||:|:|:|:|:|:|
Db 170 NKEDQRSLTRETKI 185

RESULT 3

A36469

bam protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 19-Apr-1991 #sequence_revision 19-Apr-1991 #text_change 24-Nov-1999

C:Accession: A36469

R:McKearin, D.M.; Spradling, A.C.

Genes Dev. 4, 2242-2251, 1990

A:Title: bag-of-marbles; a Drosophila gene required to initiate both male and female gametogenesis

A:Reference number: A36469; MUID:91122627

A:Accession: A36469

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-442 <MCK>

A:Cross-references: GB:X56202; NID:g2558529; PID:g7639

C:Genetics:

A:Gene: FlyBase:bam

A:Cross-references: FlyBase:FBgn0000158

C:Superfamily: fruit fly bam protein

Query Match 51.0%; Score 49; DB 2; Length 442;

Best Local Similarity 66.7%; Pred. No. 8.1; Indels 5; Gaps 0;

Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 NEREDERTLTKEYED 15

||| | | | | |

Db 35 NENEDPRKATCEYED 49

RESULT 4

T38241

probable U3 small nucleolar ribonucleoprotein component - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 07-Dec-1999

C:Accession: T38241

R:Brown, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, August 1995

A:Reference number: T21781

A:Accession: T38241

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-598 <BRO>

A:Cross-references: EMBL:X798559; PIDN:CA111156.1; GSPDB:GN00066; SPDB:SPAC23C11.03

A:Experimental source: strain 972h-; cosmid c23C11

C:Genetics:

A:Gene: SPDB:SPAC23C11.03

A:Map position: 1

Query Match 51.0%; Score 49; DB 2; Length 598;

Best Local Similarity 47.4%; Pred. No. 11;

Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 NEREDERTLTKEYEDV 19

|||:|:|:|:|:|:|

Db 401 NENKQRSIAEYEEFLK 419

RESULT 5

S61546

transferrin binding protein 2 - Neisseria meningitidis (fragment)

C:Species: Neisseria meningitidis

A:Variety: strain 2996

C:Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 26-Aug-1999

C:Accession: S61546

R:Kokbi, B.; Maitre-Wilmotte, G.; Mazarin, V.; Fourrichon, L.; Lissolo, L.; Quantin-M FEMS Microbiol. Lett. 132, 277-283, 1995

A:Title: Variable sequences in a mosaic-like domain of meningococcal tbp2 encode immunoreactive epitopes

A:Reference number: S61544; MUID:96039602

A:Accession: S61546

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-217 <ROK>

A:Cross-references: EMBL:X88867; NID:g1070368; PIDN:CAA61337.1; PID:g1070369

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995

C:Superfamily: bacterial pathogen transferrin-binding protein; tonB-dependent receptor

C:Keywords: membrane protein

Query Match 50.0%; Score 48; DB 2; Length 217;

Best Local Similarity 50.0%; Pred. No. 5.2;

Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 EREDERTLTKEYEDIVLK 19

|||:|:|:|:|:|:|

Db 183 ERTDEKEIPKEQDIVYR 200

RESULT 6

S70661

transferrin-binding protein 2 precursor - Neisseria meningitidis

C:Species: Neisseria meningitidis

C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000

C:Accession: S70661; S37626; S70659

R:Legrain, M.

submitted to the EMBL Data Library, August 1995

A:Reference number: S70661

A:Accession: S70661

A:Molecule type: DNA

A:Residues: 1-689 <LEG>

A:Cross-references: EMBL:X50732; NID:g1177570; PIDN:CAA90599.1; PID:g1177571

A:Experimental source: strain B283

R:Wilton, J.; Ala'Aldeen, D.; Palmer, H.M.; Borriello, S.P.

FEMS Microbiol. Lett. 107, 59-66, 1993

A:Title: Cloning and partial sequence of transferrin-binding protein 2 of Neisseria meningitidis

A:Reference number: S37625; MUID:93224009

A:Accession: S37626

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 25-65 <WIL>

A:Cross-references: EMBL:X75167; NID:g405780; PIDN:CAA53009.1; PID:g405781

A:Experimental source: strain B15p1.16

R:Legrain, M.; Findeli, A.; Villeval, D.; Quantin-Millet, M.J.; Jacobs, E.

Mol. Microbiol. 19, 159-169, 1996

A:Title: Molecular characterization of hybrid Tbp2 proteins from Neisseria meningitidis

A:Reference number: S70659; MUID:96419172

A:Accession: S70659

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 21-245, 'D', 247-260, 'D', 262-689 <LEW>

A:Cross-references: EMBL:X50732

C:Genetics:

A:Gene: tbp2

C:Superfamily: bacterial pathogen transferrin-binding protein; tonB-dependent receptor

C:Keywords: membrane protein

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-689/Product: transferrin-binding protein 2 #status predicted <MAT>

Query Match 50.0%; Score 48; DB 2; Length 689;

Best Local Similarity 50.0%; Pred. No. 19;

Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 EREDERTLTKEYEDIVLK 19

||| : || : ||| :

Db 526 ERTDEKEIPKEQQDIVR 543

RESULT 7

QQV215

D9R protein - vaccinia virus (strain WR)

C:Species: vaccinia virus

C>Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 16-Jul-1999

C:Accession: A03885

R:Niles, E.G.; Condit, R.C.; Caro, P.; Davidson, K.; Matusick, L.; Seto, J.

Virology 153, 96-112, 1986

A:Title: Nucleotide sequence and genetic map of the 16-kb vaccinia virus HindIII D fragm

A:Reference number: A01146; MUID:86291159

A:Accession: A03885

A:Molecule type: DNA

A:Residues: 1-213 <NILE>

A:Cross-references: GB:M15058; NID:g335640; PIDN:AAA48265.1; PID:g335653

C:Superfamily: vaccinia virus D9R protein; mutt domain homology

F:106-140/Domain: mutt domain homology <MUTT>

Query Match 47.9%; Score 46; DB 1; Length 213;

Best Local Similarity 44.4%; Pred. No. 10;

Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 EREDERTLTKEYEDIVLK 19

||| || : ||| :

Db 129 EESDERITVKEFGNVILK 146

RESULT 8

F36847

F9R protein - variola virus (strain India-1967)

N:Alternate names: D9R protein

C:Species: variola virus

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 23-Mar-2001

C:Accession: F36847; S33113

R:Blinov, V.M.

submitted to GenBank, November 1992

A:Reference number: A36859

A:Accession: F36847

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-213 <BLI>

A:Cross-references: GB:X69198; NID:g456758; PIDN:CAA49040.1; PID:g297279

R:Shchelkunov, S.N.; Blinov, V.M.; Totmenin, A.V.; Marennikova, S.S.; Kolykhalov, A.A.;

dzhaparidze, O.G.; Sandakhchiev, L.S.

Virus Res. 27, 25-35, 1993

A:Title: Nucleotide sequence analysis of variola virus HindIII M, L, I genome fragments.

A:Reference number: S33069; MUID:93190624

A:Accession: S33113

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-213 <SHC>

A:Cross-references: EMBL:X67119; NID:g62330; PIDN:CAA47598.1; PID:g62375

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1992

C:Superfamily: vaccinia virus D9R protein; mutt domain homology

F:106-140/Domain: mutt domain homology <MUTT>

Query Match 47.9%; Score 46; DB 1; Length 213;

Best Local Similarity 44.4%; Pred. No. 10;

Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 EREDERTLTKEYEDIVLK 19

||| || : ||| :

Db 129 EESDERITVKEFGNVILK 146

RESULT 9

H42515

D9R protein - vaccinia virus (strain Copenhagen)

C:Species: vaccinia virus

A>Note: host Homo sapiens (man)

C>Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 17-May-1996

C:Accession: H42515

R:Johnson, G.P.

submitted to GenBank, June 1990

A:Reference number: A33172

A:Accession: H42515

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-213 <JOH>

C:Superfamily: vaccinia virus D9R protein; mutt domain homology

F:106-140/Domain: mutt domain homology <MUTT>

Query Match 47.9%; Score 46; DB 2; Length 213;

Best Local Similarity 44.4%; Pred. No. 10;

Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 EREDERTLTKEYEDIVLK 19

||| || : ||| :

Db 129 EESDERITVKEFGNVILK 146

RESULT 10

A72163

F9R protein - variola minor virus (strain Garcia-1966)

C:Species: variola minor virus

C>Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 20-Jun-2000

C:Accession: A72163

R:Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Safonov, P.F.; Massung, R.F.; Lo

submitted to GenBank, March 1998

A:Description: Analysis of the complete coding sequence of DNA of alastrim variola m1

A:Reference number: A72150

A:Accession: A72163

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-213 <SHC>

A:Cross-references: GB:Y16780; NID:g5830555; PIDN:CAB54699.1; PID:g5830660

A:Experimental source: strain Garcia-1966

C:Genetics:

A:Gene: F9R

C:Superfamily: vaccinia virus D9R protein; mutt domain homology

Query Match 47.9%; Score 46; DB 2; Length 213;

Best Local Similarity 44.4%; Pred. No. 10;

Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 EREDERTLTKEYEDIVLK 19

||| || : ||| :

Db 129 EESDERITVKEFGNVILK 146

RESULT 11

T28537

hypothetical protein F9R - variola major virus

C:Species: variola major virus

C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000

C:Accession: T28537

R:Massung, R.F.; Esposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Au

Nature 366, 748-751, 1993

A:Title: Potential virulence determinants in terminal regions of variola smallpox vir

A:Reference number: Z20488; MUID:94088747

A:Accession: T28537

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-213 <MAS>

A:Cross-references: EMBL:L22579; NID:g623595; PIDN:AAA60847.1; PID:g439017

A:Experimental source: strain Bangladesh-1975

Mon Feb 4 15:23:42 2002

C:Superfamily: vaccinia virus D9R protein; mutT domain homology

Query Match 47.9%; Score 46; DB 2; Length 213;
 Best Local Similarity 44.4%; Pred. No. 10;
 Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 2 EREDERTLTKEYEDIVLK 19
 I : : : : :
 Db 129 EESDERITVKEFGNVILK 146

RESULT 12

T37382
 25K mutT-like protein - vaccinia virus (strain Ankara)
 C:Species: vaccinia virus
 A:Variety: strain Ankara
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 04-Mar-2000
 R:Antoine, G.; Scheiflinger, F.; Falkner, F.G.; Dorner, F.
 submitted to the EMBL Data Library, March 1997
 A:Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain
 A:Reference number: Z20877
 A:Accession: T37382
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-213 <ANT>
 A:Cross-references: EMBL:U94848; PIDN:AAB96517.1
 A:Experimental source: strain Ankara
 C:Genetics:
 A:Note: MVA106R
 C:Superfamily: vaccinia virus D9R protein; mutT domain homology

Query Match 47.9%; Score 46; DB 2; Length 213;
 Best Local Similarity 44.4%; Pred. No. 10;
 Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 2 EREDERTLTKEYEDIVLK 19
 I : : : : :
 Db 129 EESDERITVKEFGNVILK 146

RESULT 13

D84509
 hypothetical protein At2g13640 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Mar-2001
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487
 A:Accession: D84509
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-384 <STO>
 A:Cross-references: GB:AE002093; NID:g4558668; PIDN:AAD22686.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g13640
 A:Map position: 2
 C:Superfamily: Arabidopsis thaliana hypothetical protein At2g32820

Query Match 47.9%; Score 46; DB 2; Length 384;
 Best Local Similarity 57.1%; Pred. No. 20;
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NEREDERTLTKEYE 14
 I : : : : :
 Db 172 NKKEDORSITRETE 185

RESULT 14

T37768
 probable vacuolar biogenesis protein - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T37768
 R:Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, June 1997
 A:Reference number: Z21738
 A:Accession: T37768
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-860 <MUR>
 A:Cross-references: EMBL:Z97185; PIDN:CAR09996.3; GSPDB:GN00066; SPDB:SPAC16A10.03c
 A:Experimental source: strain 972h-; cosmid c16A10
 C:Genetics:
 A:Gene: SPDB:SPAC16A10.03c
 A:Map position: 1
 A:Introns: 40/3; 100/3

Query Match 47.9%; Score 46; DB 2; Length 860;
 Best Local Similarity 57.1%; Pred. No. 49;
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 6 ERTLTKYEYEDIVLK 19
 I : : : : :
 Db 703 ERSITNKYEDILYK 716

RESULT 15

T31307
 type I fatty acid synthase homolog - Cryptosporidium parvum
 C:Species: Cryptosporidium parvum
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Nov-2000
 C:Accession: T31307
 R:Zhu, G.; Marchewka, M.J.; Woods, K.M.; Upton, S.J.; Keithly, J.S.
 submitted to the EMBL Data Library, August 1998
 A:Description: Characterization of a type I FAS gene in the parasitic protozoan Crypt
 A:Reference number: Z20993
 A:Accession: T31307
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-8243 <ZHU>
 A:Cross-references: EMBL:AF082993; NID:g4092068; PID:g4092069; PIDN:AAC99407.1
 C:Genetics:
 A:Note: FAS1
 C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acetate--CoA lig
 C:Keywords: carrier protein
 F:100-714/Domain: acetate--CoA ligase homology <ACL>
 F:791-861/Domain: acyl carrier protein homology <ACP1>
 F:906-1308/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
 F:2924-2992/Domain: acyl carrier protein homology <ACP2>
 F:3062-3463/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
 F:5157-5227/Domain: acyl carrier protein homology <ACP3>
 F:5274-5689/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>

Query Match 47.9%; Score 46; DB 2; Length 8243;
 Best Local Similarity 47.4%; Pred. No. 6.2e+02;
 Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 NEREDERTLTKEYEDIVLK 19
 I : : : : :
 Db 1889 NIKEDPKIMYEYEIGLK 1907

Search completed: January 29, 2002, 10:26:47
 Job time: 658 sec

THIS PAGE BLANK (USPTO)